

BEST AVAILABLE COPY

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

**(19) World Intellectual Property Organization
International Bureau**



A standard linear barcode is located at the bottom of the page, spanning most of the width. It is used for tracking and identification of the journal issue.

**(43) International Publication Date
13 December 2001 (13.12.2001)**

PCT

(10) International Publication Number
WO 01/93983 A1

(51) International Patent Classification⁷:		B01D 39/16	60/220,664	25 July 2000 (25.07.2000)	US
(21) International Application Number:		PCT/US01/17800	60/220,666	25 July 2000 (25.07.2000)	US
			60/220,893	26 July 2000 (26.07.2000)	US
(22) International Filing Date:		1 June 2001 (01.06.2001)	PCT/US00/20710	28 July 2000 (28.07.2000)	US
			60/222,425	1 August 2000 (01.08.2000)	US
(25) Filing Language:		English	60/227,133	22 August 2000 (22.08.2000)	US
(26) Publication Language:		English	PCT/US00/23522	23 August 2000 (23.08.2000)	US
			PCT/US00/23328	24 August 2000 (24.08.2000)	US
			60/000,000	15 September 2000 (15.09.2000)	US
(30) Priority Data:			PCT/US00/30873		
PCT/US00/15264		2 June 2000 (02.06.2000)	US	10 November 2000 (10.11.2000)	US
60/209,832		5 June 2000 (05.06.2000)	US	60/253,646	28 November 2000 (28.11.2000)
60/212,901		20 June 2000 (20.06.2000)	US	PCT/US00/32678	
60/213,807		22 June 2000 (22.06.2000)	US		1 December 2000 (01.12.2000)
60/219,556		20 July 2000 (20.07.2000)	US	09/747,259	20 December 2000 (20.12.2000)
60/220,585		25 July 2000 (25.07.2000)	US	PCT/US00/34956	
60/220,605		25 July 2000 (25.07.2000)	US		20 December 2000 (20.12.2000)
60/220,607		25 July 2000 (25.07.2000)	US	PCT/US01/06520	
60/220,624		25 July 2000 (25.07.2000)	US		28 February 2001 (28.02.2001)
60/220,638		25 July 2000 (25.07.2000)	US	PCT/US01/06666	1 March 2001 (01.03.2001)

[Continued on next page]

(54) Title: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

(57) Abstract: The present invention is directed to novel polypeptides and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.



09/816,744	22 March 2001 (22.03.2001)	US
09/854,208	10 May 2001 (10.05.2001)	US
09/854,280	10 May 2001 (10.05.2001)	US
PCT/US01/17092	25 May 2001 (25.05.2001)	US

(71) **Applicant (for all designated States except US):** GENENTECH, INC. [US/US]; 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(72) **Inventors; and**

(75) **Inventors/Applicants (for US only):** BAKER, Kevin, P. [GB/US]; 14006 Indian Run Drive, Darnestown, MD 20878 (US). DESNOYERS, Luc [CA/US]; 2050 Stockton Street, San Francisco, CA 94133 (US). GERRITSEN, Mary, E. [CA/US]; 541 Parrott Drive, San Mateo, CA 94402 (US). GODDARD, Audrey [CA/US]; 110 Congo Street, San Francisco, CA 94131 (US). GODOWSKI, Paul, J. [US/US]; 2627 Easton Drive, Burlingame, CA 94010 (US). GRIMALDI, J., Christopher [US/US]; 1434 36th Avenue, San Francisco, CA 94122 (US). GURNEY, Austin, L. [US/US]; 1 Debbie Lane, Belmont, CA 94002 (US). SMITH, Victoria [AU/US]; 19 Dwight Road, Burlingame, CA 94010 (US). STEPHAN, Jean-Philippe, F. [FR/US]; 320 C Lansdale Avenue, Millbrae, CA 94030 (US). WATANABE, Colin, K. [US/US]; 128 Corliss Drive, Moraga, CA 94556 (US). WOOD, William, I. [US/US]; 35 Southdown Court, Hillsborough, CA 94010 (US).

(74) **Agents:** BARNES, Elizabeth, M. et al.; c/o Genentech, Inc., MS49, 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(81) **Designated States (national):** AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) **Designated States (regional):** ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE
SAME

FIELD OF THE INVENTION

The present invention relates generally to the identification and isolation of novel DNA and to the
5 recombinant production of novel polypeptides.

BACKGROUND OF THE INVENTION

Extracellular proteins play important roles in, among other things, the formation, differentiation and
10 maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration,
differentiation, or interaction with other cells, is typically governed by information received from other cells
and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance,
mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which
15 are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. These secreted
polypeptides or signaling molecules normally pass through the cellular secretory pathway to reach their site of
action in the extracellular environment.

Secreted proteins have various industrial applications, including as pharmaceuticals, diagnostics,
biosensors and bioreactors. Most protein drugs available at present, such as thrombolytic agents, interferons,
interleukins, erythropoietins, colony stimulating factors, and various other cytokines, are secretory proteins.
Their receptors, which are membrane proteins, also have potential as therapeutic or diagnostic agents. Efforts
20 are being undertaken by both industry and academia to identify new, native secreted proteins. Many efforts are
focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel
secreted proteins. Examples of screening methods and techniques are described in the literature [see, for
example, Klein et al., *Proc. Natl. Acad. Sci.* 93:7108-7113 (1996); U.S. Patent No. 5,536,637].

Membrane-bound proteins and receptors can play important roles in, among other things, the formation,
25 differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation,
migration, differentiation, or interaction with other cells, is typically governed by information received from other
cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for
instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and
hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins.
30 Such membrane-bound proteins and cell receptors include, but are not limited to, cytokine receptors, receptor
kinases, receptor phosphatases, receptors involved in cell-cell interactions, and cellular adhesin molecules like
selectins and integrins. For instance, transduction of signals that regulate cell growth and differentiation is
regulated in part by phosphorylation of various cellular proteins. Protein tyrosine kinases, enzymes that catalyze
that process, can also act as growth factor receptors. Examples include fibroblast growth factor receptor and

nerve growth factor receptor.

Membrane-bound proteins and receptor molecules have various industrial applications, including as pharmaceutical and diagnostic agents. Receptor immunoadhesins, for instance, can be employed as therapeutic agents to block receptor-ligand interactions. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction.

5

Efforts are being undertaken by both industry and academia to identify new, native receptor or membrane-bound proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel receptor or membrane-bound proteins.

SUMMARY OF THE INVENTION

10 In one embodiment, the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence that encodes a PRO polypeptide.

In one aspect, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule encoding a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of the full-length amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

In other aspects, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94%

nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule comprising the coding sequence of a full-length PRO polypeptide cDNA as disclosed herein, the coding sequence of a PRO polypeptide lacking the signal peptide as disclosed herein, the coding sequence of an extracellular domain of a transmembrane PRO polypeptide, with or without the signal peptide, as disclosed herein or the coding sequence of any other specifically defined fragment of the full-length amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

In a further aspect, the invention concerns an isolated nucleic acid molecule comprising a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule that encodes the same mature polypeptide encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein, or (b) the complement of the DNA molecule of (a).

Another aspect the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence encoding a PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated, or is complementary to such encoding nucleotide sequence, wherein the transmembrane domain(s) of such polypeptide are disclosed herein. Therefore, soluble extracellular domains of the herein described PRO polypeptides are contemplated.

Another embodiment is directed to fragments of a PRO polypeptide coding sequence, or the complement thereof, that may find use as, for example, hybridization probes, for encoding fragments of a PRO polypeptide that may optionally encode a polypeptide comprising a binding site for an anti-PRO antibody or as antisense oligonucleotide probes. Such nucleic acid fragments are usually at least about 10 nucleotides in length, alternatively at least about 15 nucleotides in length, alternatively at least about 20 nucleotides in length, alternatively at least about 30 nucleotides in length, alternatively at least about 40 nucleotides in length, alternatively at least about 50 nucleotides in length, alternatively at least about 60 nucleotides in length, alternatively at least about 70 nucleotides in length, alternatively at least about 80 nucleotides in length, alternatively at least about 90 nucleotides in length, alternatively at least about 100 nucleotides in length, alternatively at least about 110 nucleotides in length, alternatively at least about 120 nucleotides in length,

alternatively at least about 130 nucleotides in length, alternatively at least about 140 nucleotides in length, alternatively at least about 150 nucleotides in length, alternatively at least about 160 nucleotides in length, alternatively at least about 170 nucleotides in length, alternatively at least about 180 nucleotides in length, alternatively at least about 190 nucleotides in length, alternatively at least about 200 nucleotides in length, alternatively at least about 250 nucleotides in length, alternatively at least about 300 nucleotides in length, 5 alternatively at least about 350 nucleotides in length, alternatively at least about 400 nucleotides in length, alternatively at least about 450 nucleotides in length, alternatively at least about 500 nucleotides in length, alternatively at least about 600 nucleotides in length, alternatively at least about 700 nucleotides in length, alternatively at least about 800 nucleotides in length, alternatively at least about 900 nucleotides in length and alternatively at least about 1000 nucleotides in length, wherein in this context the term "about" means the 10 referenced nucleotide sequence length plus or minus 10% of that referenced length. It is noted that novel fragments of a PRO polypeptide-encoding nucleotide sequence may be determined in a routine manner by aligning the PRO polypeptide-encoding nucleotide sequence with other known nucleotide sequences using any of a number of well known sequence alignment programs and determining which PRO polypeptide-encoding nucleotide sequence fragment(s) are novel. All of such PRO polypeptide-encoding nucleotide sequences are contemplated 15 herein. Also contemplated are the PRO polypeptide fragments encoded by these nucleotide molecule fragments, preferably those PRO polypeptide fragments that comprise a binding site for an anti-PRO antibody.

In another embodiment, the invention provides isolated PRO polypeptide encoded by any of the isolated nucleic acid sequences hereinabove identified.

In a certain aspect, the invention concerns an isolated PRO polypeptide, comprising an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid 20 sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% 25 amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of the full-length amino acid sequence as disclosed herein.

30 In a further aspect, the invention concerns an isolated PRO polypeptide comprising an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83%

amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to an amino acid sequence encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein.

In a specific aspect, the invention provides an isolated PRO polypeptide without the N-terminal signal sequence and/or the initiating methionine and is encoded by a nucleotide sequence that encodes such an amino acid sequence as hereinbefore described. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

Another aspect the invention provides an isolated PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

In yet another embodiment, the invention concerns agonists and antagonists of a native PRO polypeptide as defined herein. In a particular embodiment, the agonist or antagonist is an anti-PRO antibody or a small molecule.

In a further embodiment, the invention concerns a method of identifying agonists or antagonists to a PRO polypeptide which comprise contacting the PRO polypeptide with a candidate molecule and monitoring a biological activity mediated by said PRO polypeptide. Preferably, the PRO polypeptide is a native PRO polypeptide.

In a still further embodiment, the invention concerns a composition of matter comprising a PRO polypeptide, or an agonist or antagonist of a PRO polypeptide as herein described, or an anti-PRO antibody, in combination with a carrier. Optionally, the carrier is a pharmaceutically acceptable carrier.

Another embodiment of the present invention is directed to the use of a PRO polypeptide, or an agonist or antagonist thereof as hereinbefore described, or an anti-PRO antibody, for the preparation of a medicament useful in the treatment of a condition which is responsive to the PRO polypeptide, an agonist or antagonist thereof or an anti-PRO antibody.

In other embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the herein described polypeptides. Host cell comprising any such vector are also provided. By way of example, the host cells may be CHO cells, *E. coli*, or yeast. A process for producing any of the herein described

polypeptides is further provided and comprises culturing host cells under conditions suitable for expression of the desired polypeptide and recovering the desired polypeptide from the cell culture.

In other embodiments, the invention provides chimeric molecules comprising any of the herein described polypeptides fused to a heterologous polypeptide or amino acid sequence. Example of such chimeric molecules comprise any of the herein described polypeptides fused to an epitope tag sequence or a Fc region of an immunoglobulin.

5 In another embodiment, the invention provides an antibody which binds, preferably specifically, to any of the above or below described polypeptides. Optionally, the antibody is a monoclonal antibody, humanized antibody, antibody fragment or single-chain antibody.

10 In yet other embodiments, the invention provides oligonucleotide probes which may be useful for isolating genomic and cDNA nucleotide sequences, measuring or detecting expression of an associated gene or as antisense probes, wherein those probes may be derived from any of the above or below described nucleotide sequences. Preferred probe lengths are described above.

15 In yet other embodiments, the present invention is directed to methods of using the PRO polypeptides of the present invention for a variety of uses based upon the functional biological assay data presented in the Examples below.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A-1B show a nucleotide sequence (SEQ ID NO:1) of a native sequence PRO6004 cDNA, wherein SEQ ID NO:1 is a clone designated herein as "DNA92259".

20 Figure 2 shows the amino acid sequence (SEQ ID NO:2) derived from the coding sequence of SEQ ID NO:1 shown in Figures 1A-1B.

Figure 3 shows a nucleotide sequence (SEQ ID NO:3) of a native sequence PRO4981 cDNA, wherein SEQ ID NO:3 is a clone designated herein as "DNA94849-2960".

25 Figure 4 shows the amino acid sequence (SEQ ID NO:4) derived from the coding sequence of SEQ ID NO:3 shown in Figure 3.

Figure 5 shows a nucleotide sequence (SEQ ID NO:5) of a native sequence PRO7174 cDNA, wherein SEQ ID NO:5 is a clone designated herein as "DNA96883-2745".

Figure 6 shows the amino acid sequence (SEQ ID NO:6) derived from the coding sequence of SEQ ID NO:5 shown in Figure 5.

30 Figure 7 shows a nucleotide sequence (SEQ ID NO:7) of a native sequence PRO5778 cDNA, wherein SEQ ID NO:7 is a clone designated herein as "DNA96894-2675".

Figure 8 shows the amino acid sequence (SEQ ID NO:8) derived from the coding sequence of SEQ ID NO:7 shown in Figure 7.

35 Figure 9 shows a nucleotide sequence (SEQ ID NO:9) of a native sequence PRO4332 cDNA, wherein SEQ ID NO:9 is a clone designated herein as "DNA100272-2969".

Figure 10 shows the amino acid sequence (SEQ ID NO:10) derived from the coding sequence of SEQ ID NO:9 shown in Figure 9.

Figure 11 shows a nucleotide sequence (SEQ ID NO:11) of a native sequence PRO9799 cDNA, wherein SEQ ID NO:11 is a clone designated herein as "DNA108696-2966".

Figure 12 shows the amino acid sequence (SEQ ID NO:12) derived from the coding sequence of SEQ ID NO:11 shown in Figure 11.

5 Figure 13 shows a nucleotide sequence (SEQ ID NO:13) of a native sequence PRO9909 cDNA, wherein SEQ ID NO:13 is a clone designated herein as "DNA117935-2801".

Figure 14 shows the amino acid sequence (SEQ ID NO:14) derived from the coding sequence of SEQ ID NO:13 shown in Figure 13.

10 Figure 15 shows a nucleotide sequence (SEQ ID NO:15) of a native sequence PRO9917 cDNA, wherein SEQ ID NO:15 is a clone designated herein as "DNA119474-2803".

Figure 16 shows the amino acid sequence (SEQ ID NO:16) derived from the coding sequence of SEQ ID NO:15 shown in Figure 15.

Figure 17 shows a nucleotide sequence (SEQ ID NO:17) of a native sequence PRO9771 cDNA, wherein SEQ ID NO:17 is a clone designated herein as "DNA119498-2965".

15 Figure 18 shows the amino acid sequence (SEQ ID NO:18) derived from the coding sequence of SEQ ID NO:17 shown in Figure 17.

Figure 19 shows a nucleotide sequence (SEQ ID NO:19) of a native sequence PRO9877 cDNA, wherein SEQ ID NO:19 is a clone designated herein as "DNA119502-2789".

Figure 20 shows the amino acid sequence (SEQ ID NO:20) derived from the coding sequence of SEQ ID NO:19 shown in Figure 19.

20 Figure 21 shows a nucleotide sequence (SEQ ID NO:21) of a native sequence PRO9903 cDNA, wherein SEQ ID NO:21 is a clone designated herein as "DNA119516-2797".

Figure 22 shows the amino acid sequence (SEQ ID NO:22) derived from the coding sequence of SEQ ID NO:21 shown in Figure 21.

25 Figure 23 shows a nucleotide sequence (SEQ ID NO:23) of a native sequence PRO9830 cDNA, wherein SEQ ID NO:23 is a clone designated herein as "DNA119530-2968".

Figure 24 shows the amino acid sequence (SEQ ID NO:24) derived from the coding sequence of SEQ ID NO:23 shown in Figure 23.

Figure 25 shows a nucleotide sequence (SEQ ID NO:25) of a native sequence PRO7155 cDNA, wherein SEQ ID NO:25 is a clone designated herein as "DNA121772-2741".

30 Figure 26 shows the amino acid sequence (SEQ ID NO:26) derived from the coding sequence of SEQ ID NO:25 shown in Figure 25.

Figure 27 shows a nucleotide sequence (SEQ ID NO:27) of a native sequence PRO9862 cDNA, wherein SEQ ID NO:27 is a clone designated herein as "DNA125148-2782".

35 Figure 28 shows the amino acid sequence (SEQ ID NO:28) derived from the coding sequence of SEQ ID NO:27 shown in Figure 27.

Figure 29 shows a nucleotide sequence (SEQ ID NO:29) of a native sequence PRO9882 cDNA, wherein SEQ ID NO:29 is a clone designated herein as "DNA125150-2793".

Figure 30 shows the amino acid sequence (SEQ ID NO:30) derived from the coding sequence of SEQ ID NO:29 shown in Figure 29.

Figure 31 shows a nucleotide sequence (SEQ ID NO:31) of a native sequence PRO9864 cDNA, wherein SEQ ID NO:31 is a clone designated herein as "DNA125151-2784".

Figure 32 shows the amino acid sequence (SEQ ID NO:32) derived from the coding sequence of SEQ 5 ID NO:31 shown in Figure 31.

Figure 33 shows a nucleotide sequence (SEQ ID NO:33) of a native sequence PRO10013 cDNA, wherein SEQ ID NO:33 is a clone designated herein as "DNA125181-2804".

Figure 34 shows the amino acid sequence (SEQ ID NO:34) derived from the coding sequence of SEQ ID NO:33 shown in Figure 33.

10 Figure 35 shows a nucleotide sequence (SEQ ID NO:35) of a native sequence PRO9885 cDNA, wherein SEQ ID NO:35 is a clone designated herein as "DNA125192-2794".

Figure 36 shows the amino acid sequence (SEQ ID NO:36) derived from the coding sequence of SEQ ID NO:35 shown in Figure 35.

15 Figure 37 shows a nucleotide sequence (SEQ ID NO:37) of a native sequence PRO9879 cDNA, wherein SEQ ID NO:37 is a clone designated herein as "DNA125196-2792".

Figure 38 shows the amino acid sequence (SEQ ID NO:38) derived from the coding sequence of SEQ ID NO:37 shown in Figure 37.

Figure 39 shows a nucleotide sequence (SEQ ID NO:39) of a native sequence PRO10111 cDNA, wherein SEQ ID NO:39 is a clone designated herein as "DNA125200-2810".

20 Figure 40 shows the amino acid sequence (SEQ ID NO:40) derived from the coding sequence of SEQ ID NO:39 shown in Figure 39.

Figure 41 shows a nucleotide sequence (SEQ ID NO:41) of a native sequence PRO9925 cDNA, wherein SEQ ID NO:41 is a clone designated herein as "DNA125214-2814".

25 Figure 42 shows the amino acid sequence (SEQ ID NO:42) derived from the coding sequence of SEQ ID NO:41 shown in Figure 41.

Figure 43 shows a nucleotide sequence (SEQ ID NO:43) of a native sequence PRO9905 cDNA, wherein SEQ ID NO:43 is a clone designated herein as "DNA125219-2799".

Figure 44 shows the amino acid sequence (SEQ ID NO:44) derived from the coding sequence of SEQ ID NO:43 shown in Figure 43.

30 Figure 45 shows a nucleotide sequence (SEQ ID NO:45) of a native sequence PRO10276 cDNA, wherein SEQ ID NO:45 is a clone designated herein as "DNA128309-2825".

Figure 46 shows the amino acid sequence (SEQ ID NO:46) derived from the coding sequence of SEQ ID NO:45 shown in Figure 45.

35 Figure 47 shows a nucleotide sequence (SEQ ID NO:47) of a native sequence PRO9898 cDNA, wherein SEQ ID NO:47 is a clone designated herein as "DNA129535-2796".

Figure 48 shows the amino acid sequence (SEQ ID NO:48) derived from the coding sequence of SEQ ID NO:47 shown in Figure 47.

Figure 49 shows a nucleotide sequence (SEQ ID NO:49) of a native sequence PRO9904 cDNA, wherein SEQ ID NO:49 is a clone designated herein as "DNA129549-2798".

Figure 50 shows the amino acid sequence (SEQ ID NO:50) derived from the coding sequence of SEQ ID NO:49 shown in Figure 49.

5 Figure 51 shows a nucleotide sequence (SEQ ID NO:51) of a native sequence PRO19632 cDNA, wherein SEQ ID NO:51 is a clone designated herein as "DNA129580-2863".

Figure 52 shows the amino acid sequence (SEQ ID NO:52) derived from the coding sequence of SEQ ID NO:51 shown in Figure 51.

Figure 53 shows a nucleotide sequence (SEQ ID NO:53) of a native sequence PRO19672 cDNA, wherein SEQ ID NO:53 is a clone designated herein as "DNA129794-2967".

10 Figure 54 shows the amino acid sequence (SEQ ID NO:54) derived from the coding sequence of SEQ ID NO:53 shown in Figure 53.

Figure 55 shows a nucleotide sequence (SEQ ID NO:55) of a native sequence PRO9783 cDNA, wherein SEQ ID NO:55 is a clone designated herein as "DNA131590-2962".

15 Figure 56 shows the amino acid sequence (SEQ ID NO:56) derived from the coding sequence of SEQ ID NO:55 shown in Figure 55.

Figure 57 shows a nucleotide sequence (SEQ ID NO:57) of a native sequence PRO10112 cDNA, wherein SEQ ID NO:57 is a clone designated herein as "DNA135173-2811".

Figure 58 shows the amino acid sequence (SEQ ID NO:58) derived from the coding sequence of SEQ ID NO:57 shown in Figure 57.

20 Figures 59A-59B show a nucleotide sequence (SEQ ID NO:59) of a native sequence PRO10284 cDNA, wherein SEQ ID NO:59 is a clone designated herein as "DNA138039-2828".

Figure 60 shows the amino acid sequence (SEQ ID NO:60) derived from the coding sequence of SEQ ID NO:59 shown in Figures 59A-59B.

25 Figure 61 shows a nucleotide sequence (SEQ ID NO:61) of a native sequence PRO10100 cDNA, wherein SEQ ID NO:61 is a clone designated herein as "DNA139540-2807".

Figure 62 shows the amino acid sequence (SEQ ID NO:62) derived from the coding sequence of SEQ ID NO:61 shown in Figure 61.

Figure 63 shows a nucleotide sequence (SEQ ID NO:63) of a native sequence PRO19628 cDNA, wherein SEQ ID NO:63 is a clone designated herein as "DNA139602-2859".

30 Figure 64 shows the amino acid sequence (SEQ ID NO:64) derived from the coding sequence of SEQ ID NO:63 shown in Figure 63.

Figure 65 shows a nucleotide sequence (SEQ ID NO:65) of a native sequence PRO19684 cDNA, wherein SEQ ID NO:65 is a clone designated herein as "DNA139632-2880".

35 Figure 66 shows the amino acid sequence (SEQ ID NO:66) derived from the coding sequence of SEQ ID NO:65 shown in Figure 65.

Figure 67 shows a nucleotide sequence (SEQ ID NO:67) of a native sequence PRO10274 cDNA, wherein SEQ ID NO:67 is a clone designated herein as "DNA139686-2823".

Figure 68 shows the amino acid sequence (SEQ ID NO:68) derived from the coding sequence of SEQ ID NO:67 shown in Figure 67.

Figure 69 shows a nucleotide sequence (SEQ ID NO:69) of a native sequence PRO9907 cDNA, wherein SEQ ID NO:69 is a clone designated herein as "DNA142392-2800".

Figure 70 shows the amino acid sequence (SEQ ID NO:70) derived from the coding sequence of SEQ 5 ID NO:69 shown in Figure 69.

Figure 71 shows a nucleotide sequence (SEQ ID NO:71) of a native sequence PRO9873 cDNA, wherein SEQ ID NO:71 is a clone designated herein as "DNA143076-2787".

Figure 72 shows the amino acid sequence (SEQ ID NO:72) derived from the coding sequence of SEQ ID NO:71 shown in Figure 71.

10 Figure 73 shows a nucleotide sequence (SEQ ID NO:73) of a native sequence PRO10201 cDNA, wherein SEQ ID NO:73 is a clone designated herein as "DNA143294-2818".

Figure 74 shows the amino acid sequence (SEQ ID NO:74) derived from the coding sequence of SEQ ID NO:73 shown in Figure 73.

15 Figure 75 shows a nucleotide sequence (SEQ ID NO:75) of a native sequence PRO10200 cDNA, wherein SEQ ID NO:75 is a clone designated herein as "DNA143514-2817".

Figure 76 shows the amino acid sequence (SEQ ID NO:76) derived from the coding sequence of SEQ ID NO:75 shown in Figure 75.

Figure 77 shows a nucleotide sequence (SEQ ID NO:77) of a native sequence PRO10196 cDNA, wherein SEQ ID NO:77 is a clone designated herein as "DNA144841-2816".

20 Figure 78 shows the amino acid sequence (SEQ ID NO:78) derived from the coding sequence of SEQ ID NO:77 shown in Figure 77.

Figure 79 shows a nucleotide sequence (SEQ ID NO:79) of a native sequence PRO10282 cDNA, wherein SEQ ID NO:79 is a clone designated herein as "DNA148380-2827".

25 Figure 80 shows the amino acid sequence (SEQ ID NO:80) derived from the coding sequence of SEQ ID NO:79 shown in Figure 79.

Figure 81 shows a nucleotide sequence (SEQ ID NO:81) of a native sequence PRO19650 cDNA, wherein SEQ ID NO:81 is a clone designated herein as "DNA149995-2871".

Figure 82 shows the amino acid sequence (SEQ ID NO:82) derived from the coding sequence of SEQ ID NO:81 shown in Figure 81.

30 Figure 83 shows a nucleotide sequence (SEQ ID NO:83) of a native sequence PRO21184 cDNA, wherein SEQ ID NO:83 is a clone designated herein as "DNA167678-2963".

Figure 84 shows the amino acid sequence (SEQ ID NO:84) derived from the coding sequence of SEQ ID NO:83 shown in Figure 83.

35 Figure 85 shows a nucleotide sequence (SEQ ID NO:85) of a native sequence PRO21201 cDNA, wherein SEQ ID NO:85 is a clone designated herein as "DNA168028-2956".

Figure 86 shows the amino acid sequence (SEQ ID NO:86) derived from the coding sequence of SEQ ID NO:85 shown in Figure 85.

Figure 87 shows a nucleotide sequence (SEQ ID NO:87) of a native sequence PRO21175 cDNA, wherein SEQ ID NO:87 is a clone designated herein as "DNA173894-2947".

Figure 88 shows the amino acid sequence (SEQ ID NO:88) derived from the coding sequence of SEQ ID NO:87 shown in Figure 87.

5 Figure 89 shows a nucleotide sequence (SEQ ID NO:89) of a native sequence PRO21340 cDNA, wherein SEQ ID NO:89 is a clone designated herein as "DNA176775-2957".

Figure 90 shows the amino acid sequence (SEQ ID NO:90) derived from the coding sequence of SEQ ID NO:89 shown in Figure 89.

Figure 91 shows a nucleotide sequence (SEQ ID NO:91) of a native sequence PRO21384 cDNA, wherein SEQ ID NO:91 is a clone designated herein as "DNA177313-2982".

10 Figure 92 shows the amino acid sequence (SEQ ID NO:92) derived from the coding sequence of SEQ ID NO:91 shown in Figure 91.

Figure 93 shows a nucleotide sequence (SEQ ID NO:93) of a native sequence PRO982 cDNA, wherein SEQ ID NO:93 is a clone designated herein as "DNA57700-1408".

15 Figure 94 shows the amino acid sequence (SEQ ID NO:94) derived from the coding sequence of SEQ ID NO:93 shown in Figure 93.

Figure 95 shows a nucleotide sequence (SEQ ID NO:95) of a native sequence PRO1160 cDNA, wherein SEQ ID NO:95 is a clone designated herein as "DNA62872-1509".

Figure 96 shows the amino acid sequence (SEQ ID NO:96) derived from the coding sequence of SEQ ID NO:95 shown in Figure 95.

20 Figure 97 shows a nucleotide sequence (SEQ ID NO:97) of a native sequence PRO1187 cDNA, wherein SEQ ID NO:97 is a clone designated herein as "DNA62876-1517".

Figure 98 shows the amino acid sequence (SEQ ID NO:98) derived from the coding sequence of SEQ ID NO:97 shown in Figure 97.

25 Figure 99 shows a nucleotide sequence (SEQ ID NO:99) of a native sequence PRO1329 cDNA, wherein SEQ ID NO:99 is a clone designated herein as "DNA66660-1585".

Figure 100 shows the amino acid sequence (SEQ ID NO:100) derived from the coding sequence of SEQ ID NO:99 shown in Figure 99.

Figure 101 shows a nucleotide sequence (SEQ ID NO:101) of a native sequence PRO231 cDNA, wherein SEQ ID NO:101 is a clone designated herein as "DNA34434-1139".

30 Figure 102 shows the amino acid sequence (SEQ ID NO:102) derived from the coding sequence of SEQ ID NO:101 shown in Figure 101.

Figure 103 shows a nucleotide sequence (SEQ ID NO:103) of a native sequence PRO357 cDNA, wherein SEQ ID NO:103 is a clone designated herein as "DNA44804-1248".

35 Figure 104 shows the amino acid sequence (SEQ ID NO:104) derived from the coding sequence of SEQ ID NO:103 shown in Figure 103.

Figure 105 shows a nucleotide sequence (SEQ ID NO:105) of a native sequence PRO725 cDNA, wherein SEQ ID NO:105 is a clone designated herein as "DNA52758-1399".

Figure 106 shows the amino acid sequence (SEQ ID NO:106) derived from the coding sequence of SEQ ID NO:105 shown in Figure 105.

Figure 107 shows a nucleotide sequence (SEQ ID NO:107) of a native sequence PRO1155 cDNA, wherein SEQ ID NO:107 is a clone designated herein as "DNA59849-1504".

Figure 108 shows the amino acid sequence (SEQ ID NO:108) derived from the coding sequence of SEQ 5 ID NO:107 shown in Figure 107.

Figure 109 shows a nucleotide sequence (SEQ ID NO:109) of a native sequence PRO1306 cDNA, wherein SEQ ID NO:109 is a clone designated herein as "DNA65410-1569".

Figure 110 shows the amino acid sequence (SEQ ID NO:110) derived from the coding sequence of SEQ ID NO:109 shown in Figure 109.

10 Figure 111 shows a nucleotide sequence (SEQ ID NO:111) of a native sequence PRO1419 cDNA, wherein SEQ ID NO:111 is a clone designated herein as "DNA71290-1630".

Figure 112 shows the amino acid sequence (SEQ ID NO:112) derived from the coding sequence of SEQ ID NO:111 shown in Figure 111.

15 Figure 113 shows a nucleotide sequence (SEQ ID NO:113) of a native sequence PRO229 cDNA, wherein SEQ ID NO:113 is a clone designated herein as "DNA33100-1159".

Figure 114 shows the amino acid sequence (SEQ ID NO:114) derived from the coding sequence of SEQ ID NO:113 shown in Figure 113.

Figure 115 shows a nucleotide sequence (SEQ ID NO:115) of a native sequence PRO1272 cDNA, wherein SEQ ID NO:115 is a clone designated herein as "DNA64896-1539".

20 Figure 116 shows the amino acid sequence (SEQ ID NO:116) derived from the coding sequence of SEQ ID NO:115 shown in Figure 115.

Figure 117 shows a nucleotide sequence (SEQ ID NO:117) of a native sequence PRO4405 cDNA, wherein SEQ ID NO:117 is a clone designated herein as "DNA84920-2614".

25 Figure 118 shows the amino acid sequence (SEQ ID NO:118) derived from the coding sequence of SEQ ID NO:117 shown in Figure 117.

Figure 119 shows a nucleotide sequence (SEQ ID NO:119) of a native sequence PRO181 cDNA, wherein SEQ ID NO:119 is a clone designated herein as "DNA23330-1390".

Figure 120 shows the amino acid sequence (SEQ ID NO:120) derived from the coding sequence of SEQ ID NO:119 shown in Figure 119.

30 Figure 121 shows a nucleotide sequence (SEQ ID NO:121) of a native sequence PRO214 cDNA, wherein SEQ ID NO:121 is a clone designated herein as "DNA32286-1191".

Figure 122 shows the amino acid sequence (SEQ ID NO:122) derived from the coding sequence of SEQ ID NO:121 shown in Figure 121.

35 Figure 123 shows a nucleotide sequence (SEQ ID NO:123) of a native sequence PRO247 cDNA, wherein SEQ ID NO:123 is a clone designated herein as "DNA35673-1201".

Figure 124 shows the amino acid sequence (SEQ ID NO:124) derived from the coding sequence of SEQ ID NO:123 shown in Figure 123.

Figure 125 shows a nucleotide sequence (SEQ ID NO:125) of a native sequence PRO337 cDNA, wherein SEQ ID NO:125 is a clone designated herein as "DNA43316-1237".

Figure 126 shows the amino acid sequence (SEQ ID NO:126) derived from the coding sequence of SEQ ID NO:125 shown in Figure 125.

5 Figure 127 shows a nucleotide sequence (SEQ ID NO:127) of a native sequence PRO526 cDNA, wherein SEQ ID NO:127 is a clone designated herein as "DNA44184-1319".

Figure 128 shows the amino acid sequence (SEQ ID NO:128) derived from the coding sequence of SEQ ID NO:127 shown in Figure 127.

Figure 129 shows a nucleotide sequence (SEQ ID NO:129) of a native sequence PRO363 cDNA, wherein SEQ ID NO:129 is a clone designated herein as "DNA45419-1252".

10 Figure 130 shows the amino acid sequence (SEQ ID NO:130) derived from the coding sequence of SEQ ID NO:129 shown in Figure 129.

Figure 131 shows a nucleotide sequence (SEQ ID NO:131) of a native sequence PRO531 cDNA, wherein SEQ ID NO:131 is a clone designated herein as "DNA48314-1320".

15 Figure 132 shows the amino acid sequence (SEQ ID NO:132) derived from the coding sequence of SEQ ID NO:131 shown in Figure 131.

Figure 133 shows a nucleotide sequence (SEQ ID NO:133) of a native sequence PRO1083 cDNA, wherein SEQ ID NO:133 is a clone designated herein as "DNA50921-1458".

Figure 134 shows the amino acid sequence (SEQ ID NO:134) derived from the coding sequence of SEQ ID NO:133 shown in Figure 133.

20 Figure 135 shows a nucleotide sequence (SEQ ID NO:135) of a native sequence PRO840 cDNA, wherein SEQ ID NO:135 is a clone designated herein as "DNA53987".

Figure 136 shows the amino acid sequence (SEQ ID NO:136) derived from the coding sequence of SEQ ID NO:135 shown in Figure 135.

25 Figure 137 shows a nucleotide sequence (SEQ ID NO:137) of a native sequence PRO1080 cDNA, wherein SEQ ID NO:137 is a clone designated herein as "DNA56047-1456".

Figure 138 shows the amino acid sequence (SEQ ID NO:138) derived from the coding sequence of SEQ ID NO:137 shown in Figure 137.

Figure 139 shows a nucleotide sequence (SEQ ID NO:139) of a native sequence PRO788 cDNA, wherein SEQ ID NO:139 is a clone designated herein as "DNA56405-1357".

30 Figure 140 shows the amino acid sequence (SEQ ID NO:140) derived from the coding sequence of SEQ ID NO:139 shown in Figure 139.

Figure 141 shows a nucleotide sequence (SEQ ID NO:141) of a native sequence PRO1478 cDNA, wherein SEQ ID NO:141 is a clone designated herein as "DNA56531-1648".

35 Figure 142 shows the amino acid sequence (SEQ ID NO:142) derived from the coding sequence of SEQ ID NO:141 shown in Figure 141.

Figure 143 shows a nucleotide sequence (SEQ ID NO:143) of a native sequence PRO1134 cDNA, wherein SEQ ID NO:143 is a clone designated herein as "DNA56865-1491".

Figure 144 shows the amino acid sequence (SEQ ID NO:144) derived from the coding sequence of SEQ ID NO:143 shown in Figure 143.

Figure 145 shows a nucleotide sequence (SEQ ID NO:145) of a native sequence PRO826 cDNA, wherein SEQ ID NO:145 is a clone designated herein as "DNA57694-1341".

5 Figure 146 shows the amino acid sequence (SEQ ID NO:146) derived from the coding sequence of SEQ ID NO:145 shown in Figure 145.

Figure 147 shows a nucleotide sequence (SEQ ID NO:147) of a native sequence PRO1005 cDNA, wherein SEQ ID NO:147 is a clone designated herein as "DNA57708-1411".

Figure 148 shows the amino acid sequence (SEQ ID NO:148) derived from the coding sequence of SEQ ID NO:147 shown in Figure 147.

10 Figure 149 shows a nucleotide sequence (SEQ ID NO:149) of a native sequence PRO809 cDNA, wherein SEQ ID NO:149 is a clone designated herein as "DNA57836-1338".

Figure 150 shows the amino acid sequence (SEQ ID NO:150) derived from the coding sequence of SEQ ID NO:149 shown in Figure 149.

15 Figure 151 shows a nucleotide sequence (SEQ ID NO:151) of a native sequence PRO1194 cDNA, wherein SEQ ID NO:151 is a clone designated herein as "DNA57841-1522".

Figure 152 shows the amino acid sequence (SEQ ID NO:152) derived from the coding sequence of SEQ ID NO:151 shown in Figure 151.

Figure 153 shows a nucleotide sequence (SEQ ID NO:153) of a native sequence PRO1071 cDNA, wherein SEQ ID NO:153 is a clone designated herein as "DNA58847-1383".

20 Figure 154 shows the amino acid sequence (SEQ ID NO:154) derived from the coding sequence of SEQ ID NO:153 shown in Figure 153.

Figure 155 shows a nucleotide sequence (SEQ ID NO:155) of a native sequence PRO1411 cDNA, wherein SEQ ID NO:155 is a clone designated herein as "DNA59212-1627".

25 Figure 156 shows the amino acid sequence (SEQ ID NO:156) derived from the coding sequence of SEQ ID NO:155 shown in Figure 155.

Figure 157 shows a nucleotide sequence (SEQ ID NO:157) of a native sequence PRO1309 cDNA, wherein SEQ ID NO:157 is a clone designated herein as "DNA59588-1571".

Figure 158 shows the amino acid sequence (SEQ ID NO:158) derived from the coding sequence of SEQ ID NO:157 shown in Figure 157.

30 Figure 159 shows a nucleotide sequence (SEQ ID NO:159) of a native sequence PRO1025 cDNA, wherein SEQ ID NO:159 is a clone designated herein as "DNA59622-1334".

Figure 160 shows the amino acid sequence (SEQ ID NO:160) derived from the coding sequence of SEQ ID NO:159 shown in Figure 159.

35 Figure 161 shows a nucleotide sequence (SEQ ID NO:161) of a native sequence PRO1181 cDNA, wherein SEQ ID NO:161 is a clone designated herein as "DNA59847-2510".

Figure 162 shows the amino acid sequence (SEQ ID NO:162) derived from the coding sequence of SEQ ID NO:161 shown in Figure 161.

Figure 163 shows a nucleotide sequence (SEQ ID NO:163) of a native sequence PRO1126 cDNA, wherein SEQ ID NO:163 is a clone designated herein as "DNA60615-1483".

Figure 164 shows the amino acid sequence (SEQ ID NO:164) derived from the coding sequence of SEQ ID NO:163 shown in Figure 163.

Figure 165 shows a nucleotide sequence (SEQ ID NO:165) of a native sequence PRO1186 cDNA, 5 wherein SEQ ID NO:165 is a clone designated herein as "DNA60621-1516".

Figure 166 shows the amino acid sequence (SEQ ID NO:166) derived from the coding sequence of SEQ ID NO:165 shown in Figure 165.

Figure 167 shows a nucleotide sequence (SEQ ID NO:167) of a native sequence PRO1192 cDNA, wherein SEQ ID NO:167 is a clone designated herein as "DNA62814-1521".

10 Figure 168 shows the amino acid sequence (SEQ ID NO:168) derived from the coding sequence of SEQ ID NO:167 shown in Figure 167.

Figure 169 shows a nucleotide sequence (SEQ ID NO:169) of a native sequence PRO1244 cDNA, wherein SEQ ID NO:169 is a clone designated herein as "DNA64883-1526".

15 Figure 170 shows the amino acid sequence (SEQ ID NO:170) derived from the coding sequence of SEQ ID NO:169 shown in Figure 169.

Figure 171 shows a nucleotide sequence (SEQ ID NO:171) of a native sequence PRO1274 cDNA, wherein SEQ ID NO:171 is a clone designated herein as "DNA64889-1541".

Figure 172 shows the amino acid sequence (SEQ ID NO:172) derived from the coding sequence of SEQ ID NO:171 shown in Figure 171.

20 Figure 173 shows a nucleotide sequence (SEQ ID NO:173) of a native sequence PRO1412 cDNA, wherein SEQ ID NO:173 is a clone designated herein as "DNA64897-1628".

Figure 174 shows the amino acid sequence (SEQ ID NO:174) derived from the coding sequence of SEQ ID NO:173 shown in Figure 173.

25 Figure 175 shows a nucleotide sequence (SEQ ID NO:175) of a native sequence PRO1286 cDNA, wherein SEQ ID NO:175 is a clone designated herein as "DNA64903-1553".

Figure 176 shows the amino acid sequence (SEQ ID NO:176) derived from the coding sequence of SEQ ID NO:175 shown in Figure 175.

Figure 177 shows a nucleotide sequence (SEQ ID NO:177) of a native sequence PRO1330 cDNA, wherein SEQ ID NO:177 is a clone designated herein as "DNA64907-1163-1".

30 Figure 178 shows the amino acid sequence (SEQ ID NO:178) derived from the coding sequence of SEQ ID NO:177 shown in Figure 177.

Figure 179 shows a nucleotide sequence (SEQ ID NO:179) of a native sequence PRO1347 cDNA, wherein SEQ ID NO:179 is a clone designated herein as "DNA64950-1590".

35 Figure 180 shows the amino acid sequence (SEQ ID NO:180) derived from the coding sequence of SEQ ID NO:179 shown in Figure 179.

Figure 181 shows a nucleotide sequence (SEQ ID NO:181) of a native sequence PRO1305 cDNA, wherein SEQ ID NO:181 is a clone designated herein as "DNA64952-1568".

Figure 182 shows the amino acid sequence (SEQ ID NO:182) derived from the coding sequence of SEQ ID NO:181 shown in Figure 181.

Figure 183 shows a nucleotide sequence (SEQ ID NO:183) of a native sequence PRO1273 cDNA, wherein SEQ ID NO:183 is a clone designated herein as "DNA65402-1540".

5 Figure 184 shows the amino acid sequence (SEQ ID NO:184) derived from the coding sequence of SEQ ID NO:183 shown in Figure 183.

Figure 185 shows a nucleotide sequence (SEQ ID NO:185) of a native sequence PRO1279 cDNA, wherein SEQ ID NO:185 is a clone designated herein as "DNA65405-1547".

Figure 186 shows the amino acid sequence (SEQ ID NO:186) derived from the coding sequence of SEQ ID NO:185 shown in Figure 185.

10 Figure 187 shows a nucleotide sequence (SEQ ID NO:187) of a native sequence PRO1340 cDNA, wherein SEQ ID NO:187 is a clone designated herein as "DNA66663-1598".

Figure 188 shows the amino acid sequence (SEQ ID NO:188) derived from the coding sequence of SEQ ID NO:187 shown in Figure 187.

15 Figure 189 shows a nucleotide sequence (SEQ ID NO:189) of a native sequence PRO1338 cDNA, wherein SEQ ID NO:189 is a clone designated herein as "DNA66667".

Figure 190 shows the amino acid sequence (SEQ ID NO:190) derived from the coding sequence of SEQ ID NO:189 shown in Figure 189.

Figure 191 shows a nucleotide sequence (SEQ ID NO:191) of a native sequence PRO1343 cDNA, wherein SEQ ID NO:191 is a clone designated herein as "DNA66675-1587".

20 Figure 192 shows the amino acid sequence (SEQ ID NO:192) derived from the coding sequence of SEQ ID NO:191 shown in Figure 191.

Figure 193 shows a nucleotide sequence (SEQ ID NO:193) of a native sequence PRO1376 cDNA, wherein SEQ ID NO:193 is a clone designated herein as "DNA67300-1605".

25 Figure 194 shows the amino acid sequence (SEQ ID NO:194) derived from the coding sequence of SEQ ID NO:193 shown in Figure 193.

Figure 195 shows a nucleotide sequence (SEQ ID NO:195) of a native sequence PRO1387 cDNA, wherein SEQ ID NO:195 is a clone designated herein as "DNA68872-1620".

Figure 196 shows the amino acid sequence (SEQ ID NO:196) derived from the coding sequence of SEQ ID NO:195 shown in Figure 195.

30 Figure 197 shows a nucleotide sequence (SEQ ID NO:197) of a native sequence PRO1409 cDNA, wherein SEQ ID NO:197 is a clone designated herein as "DNA71269-1621".

Figure 198 shows the amino acid sequence (SEQ ID NO:198) derived from the coding sequence of SEQ ID NO:197 shown in Figure 197.

Figure 199 shows a nucleotide sequence (SEQ ID NO:199) of a native sequence PRO1488 cDNA, wherein SEQ ID NO:199 is a clone designated herein as "DNA73736-1657".

35 Figure 200 shows the amino acid sequence (SEQ ID NO:200) derived from the coding sequence of SEQ ID NO:199 shown in Figure 199.

Figure 201 shows a nucleotide sequence (SEQ ID NO:201) of a native sequence PRO1474 cDNA, wherein SEQ ID NO:201 is a clone designated herein as "DNA73739-1645".

Figure 202 shows the amino acid sequence (SEQ ID NO:202) derived from the coding sequence of SEQ ID NO:201 shown in Figure 201.

5 Figure 203 shows a nucleotide sequence (SEQ ID NO:203) of a native sequence PRO1917 cDNA, wherein SEQ ID NO:203 is a clone designated herein as "DNA76400-2528".

Figure 204 shows the amino acid sequence (SEQ ID NO:204) derived from the coding sequence of SEQ ID NO:203 shown in Figure 203.

Figure 205 shows a nucleotide sequence (SEQ ID NO:205) of a native sequence PRO1760 cDNA, wherein SEQ ID NO:205 is a clone designated herein as "DNA76532-1702".

10 Figure 206 shows the amino acid sequence (SEQ ID NO:206) derived from the coding sequence of SEQ ID NO:205 shown in Figure 205.

Figure 207 shows a nucleotide sequence (SEQ ID NO:207) of a native sequence PRO1567 cDNA, wherein SEQ ID NO:207 is a clone designated herein as "DNA76541-1675".

15 Figure 208 shows the amino acid sequence (SEQ ID NO:208) derived from the coding sequence of SEQ ID NO:207 shown in Figure 207.

Figure 209 shows a nucleotide sequence (SEQ ID NO:209) of a native sequence PRO1887 cDNA, wherein SEQ ID NO:209 is a clone designated herein as "DNA79862-2522".

Figure 210 shows the amino acid sequence (SEQ ID NO:210) derived from the coding sequence of SEQ ID NO:209 shown in Figure 209.

20 Figure 211 shows a nucleotide sequence (SEQ ID NO:211) of a native sequence PRO1928 cDNA, wherein SEQ ID NO:211 is a clone designated herein as "DNA81754-2532".

Figure 212 shows the amino acid sequence (SEQ ID NO:212) derived from the coding sequence of SEQ ID NO:211 shown in Figure 211.

25 Figure 213 shows a nucleotide sequence (SEQ ID NO:213) of a native sequence PRO4341 cDNA, wherein SEQ ID NO:213 is a clone designated herein as "DNA81761-2583".

Figure 214 shows the amino acid sequence (SEQ ID NO:214) derived from the coding sequence of SEQ ID NO:213 shown in Figure 213.

Figure 215 shows a nucleotide sequence (SEQ ID NO:215) of a native sequence PRO5723 cDNA, wherein SEQ ID NO:215 is a clone designated herein as "DNA82361".

30 Figure 216 shows the amino acid sequence (SEQ ID NO:216) derived from the coding sequence of SEQ ID NO:215 shown in Figure 215.

Figure 217 shows a nucleotide sequence (SEQ ID NO:217) of a native sequence PRO1801 cDNA, wherein SEQ ID NO:217 is a clone designated herein as "DNA83500-2506".

35 Figure 218 shows the amino acid sequence (SEQ ID NO:218) derived from the coding sequence of SEQ ID NO:217 shown in Figure 217.

Figure 219 shows a nucleotide sequence (SEQ ID NO:219) of a native sequence PRO4333 cDNA, wherein SEQ ID NO:219 is a clone designated herein as "DNA84210-2576".

Figure 220 shows the amino acid sequence (SEQ ID NO:220) derived from the coding sequence of SEQ ID NO:219 shown in Figure 219.

Figure 221 shows a nucleotide sequence (SEQ ID NO:221) of a native sequence PRO3543 cDNA, wherein SEQ ID NO:221 is a clone designated herein as "DNA86571-2551".

Figure 222 shows the amino acid sequence (SEQ ID NO:222) derived from the coding sequence of SEQ ID NO:221 shown in Figure 221.

Figure 223 shows a nucleotide sequence (SEQ ID NO:223) of a native sequence PRO3444 cDNA, wherein SEQ ID NO:223 is a clone designated herein as "DNA87997".

Figure 224 shows the amino acid sequence (SEQ ID NO:224) derived from the coding sequence of SEQ ID NO:223 shown in Figure 223.

Figure 225 shows a nucleotide sequence (SEQ ID NO:225) of a native sequence PRO4302 cDNA, wherein SEQ ID NO:225 is a clone designated herein as "DNA92218-2554".

Figure 226 shows the amino acid sequence (SEQ ID NO:226) derived from the coding sequence of SEQ ID NO:225 shown in Figure 225.

Figure 227 shows a nucleotide sequence (SEQ ID NO:227) of a native sequence PRO4322 cDNA, wherein SEQ ID NO:227 is a clone designated herein as "DNA92223-2567".

Figure 228 shows the amino acid sequence (SEQ ID NO:228) derived from the coding sequence of SEQ ID NO:227 shown in Figure 227.

Figure 229 shows a nucleotide sequence (SEQ ID NO:229) of a native sequence PRO5725 cDNA, wherein SEQ ID NO:229 is a clone designated herein as "DNA92265-2669".

Figure 230 shows the amino acid sequence (SEQ ID NO:230) derived from the coding sequence of SEQ ID NO:229 shown in Figure 229.

Figure 231 shows a nucleotide sequence (SEQ ID NO:231) of a native sequence PRO4408 cDNA, wherein SEQ ID NO:231 is a clone designated herein as "DNA92274-2617".

Figure 232 shows the amino acid sequence (SEQ ID NO:232) derived from the coding sequence of SEQ ID NO:231 shown in Figure 231.

Figure 233 shows a nucleotide sequence (SEQ ID NO:233) of a native sequence PRO9940 cDNA, wherein SEQ ID NO:233 is a clone designated herein as "DNA92282".

Figure 234 shows the amino acid sequence (SEQ ID NO:234) derived from the coding sequence of SEQ ID NO:233 shown in Figure 233.

Figure 235 shows a nucleotide sequence (SEQ ID NO:235) of a native sequence PRO7154 cDNA, wherein SEQ ID NO:235 is a clone designated herein as "DNA108760-2740".

Figure 236 shows the amino acid sequence (SEQ ID NO:236) derived from the coding sequence of SEQ ID NO:235 shown in Figure 235.

Figure 237 shows a nucleotide sequence (SEQ ID NO:237) of a native sequence PRO7425 cDNA, wherein SEQ ID NO:237 is a clone designated herein as "DNA108792-2753".

Figure 238 shows the amino acid sequence (SEQ ID NO:238) derived from the coding sequence of SEQ ID NO:237 shown in Figure 237.

Figure 239 shows a nucleotide sequence (SEQ ID NO:239) of a native sequence PRO6079 cDNA, wherein SEQ ID NO:239 is a clone designated herein as "DNA111750-2706".

Figure 240 shows the amino acid sequence (SEQ ID NO:240) derived from the coding sequence of SEQ ID NO:239 shown in Figure 239.

5 Figure 241 shows a nucleotide sequence (SEQ ID NO:241) of a native sequence PRO9836 cDNA, wherein SEQ ID NO:241 is a clone designated herein as "DNA119514-2772".

Figure 242 shows the amino acid sequence (SEQ ID NO:242) derived from the coding sequence of SEQ ID NO:241 shown in Figure 241.

Figure 243 shows a nucleotide sequence (SEQ ID NO:243) of a native sequence PRO10096 cDNA, wherein SEQ ID NO:243 is a clone designated herein as "DNA125185-2806".

10 Figure 244 shows the amino acid sequence (SEQ ID NO:244) derived from the coding sequence of SEQ ID NO:243 shown in Figure 243.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

I. Definitions

15 The terms "PRO polypeptide" and "PRO" as used herein and when immediately followed by a numerical designation refer to various polypeptides, wherein the complete designation (i.e., PRO/number) refers to specific polypeptide sequences as described herein. The terms "PRO/number polypeptide" and "PRO/number" wherein the term "number" is provided as an actual numerical designation as used herein encompass native sequence polypeptides and polypeptide variants (which are further defined herein). The PRO polypeptides described herein 20 may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods. The term "PRO polypeptide" refers to each individual PRO/number polypeptide disclosed herein. All disclosures in this specification which refer to the "PRO polypeptide" refer to each of the polypeptides individually as well as jointly. For example, descriptions of the preparation of, purification of, derivation of, formation of antibodies to or against, administration of, compositions containing, 25 treatment of a disease with, etc., pertain to each polypeptide of the invention individually. The term "PRO polypeptide" also includes variants of the PRO/number polypeptides disclosed herein.

A "native sequence PRO polypeptide" comprises a polypeptide having the same amino acid sequence as the corresponding PRO polypeptide derived from nature. Such native sequence PRO polypeptides can be isolated from nature or can be produced by recombinant or synthetic means. The term "native sequence PRO polypeptide" 30 specifically encompasses naturally-occurring truncated or secreted forms of the specific PRO polypeptide (e.g., an extracellular domain sequence), naturally-occurring variant forms (e.g., alternatively spliced forms) and naturally-occurring allelic variants of the polypeptide. In various embodiments of the invention, the native sequence PRO polypeptides disclosed herein are mature or full-length native sequence polypeptides comprising the full-length amino acids sequences shown in the accompanying figures. Start and stop codons are shown in bold font and underlined in the figures. However, while the PRO polypeptide disclosed in the accompanying figures are shown to begin with methionine residues designated herein as amino acid position 1 in the figures, it is conceivable and possible that other methionine residues located either upstream or downstream from the amino 35

acid position 1 in the figures may be employed as the starting amino acid residue for the PRO polypeptides.

The PRO polypeptide "extracellular domain" or "ECD" refers to a form of the PRO polypeptide which is essentially free of the transmembrane and cytoplasmic domains. Ordinarily, a PRO polypeptide ECD will have less than 1% of such transmembrane and/or cytoplasmic domains and preferably, will have less than 0.5% of such domains. It will be understood that any transmembrane domains identified for the PRO polypeptides of the 5 present invention are identified pursuant to criteria routinely employed in the art for identifying that type of hydrophobic domain. The exact boundaries of a transmembrane domain may vary but most likely by no more than about 5 amino acids at either end of the domain as initially identified herein. Optionally, therefore, an extracellular domain of a PRO polypeptide may contain from about 5 or fewer amino acids on either side of the transmembrane domain/extracellular domain boundary as identified in the Examples or specification and such 10 polypeptides, with or without the associated signal peptide, and nucleic acid encoding them, are contemplated by the present invention.

The approximate location of the "signal peptides" of the various PRO polypeptides disclosed herein are shown in the present specification and/or the accompanying figures. It is noted, however, that the C-terminal boundary of a signal peptide may vary, but most likely by no more than about 5 amino acids on either side of the 15 signal peptide C-terminal boundary as initially identified herein, wherein the C-terminal boundary of the signal peptide may be identified pursuant to criteria routinely employed in the art for identifying that type of amino acid sequence element (e.g., Nielsen et al., Prot. Eng. 10:1-6 (1997) and von Heinje et al., Nucl. Acids. Res. 14:4683-4690 (1986)). Moreover, it is also recognized that, in some cases, cleavage of a signal sequence from a secreted polypeptide is not entirely uniform, resulting in more than one secreted species. These mature 20 polypeptides, where the signal peptide is cleaved within no more than about 5 amino acids on either side of the C-terminal boundary of the signal peptide as identified herein, and the polynucleotides encoding them, are contemplated by the present invention.

"PRO polypeptide variant" means an active PRO polypeptide as defined above or below having at least about 80% amino acid sequence identity with a full-length native sequence PRO polypeptide sequence as disclosed 25 herein, a PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Such PRO polypeptide variants include, for instance, PRO polypeptides wherein one or more amino acid residues are added, or deleted, at the N- or C-terminus of the full-length native amino acid sequence. Ordinarily, a PRO polypeptide variant will have at least about 80% amino 30 acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least 35 about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid

sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to a full-length native sequence PRO polypeptide sequence as disclosed herein, a PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, PRO variant polypeptides are at least about 10 amino acids in length, alternatively at least about 20 amino acids in length, alternatively at least about 30 amino acids in length, alternatively at least about 40 amino acids in length, alternatively at least about 50 amino acids in length, alternatively at least about 60 amino acids in length, alternatively at least about 70 amino acids in length, alternatively at least about 80 amino acids in length, alternatively at least about 90 amino acids in length, alternatively at least about 100 amino acids in length, alternatively at least about 150 amino acids in length, alternatively at least about 200 amino acids in length, alternatively at least about 300 amino acids in length, or more.

"Percent (%) amino acid sequence identity" with respect to the PRO polypeptide sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the specific PRO polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. For purposes herein, however, % amino acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

In situations where ALIGN-2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

35

100 times the fraction X/Y

where X is the number of amino acid residues scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A. As examples of % amino acid sequence identity calculations using this method, Tables 2 and 3 5 demonstrate how to calculate the % amino acid sequence identity of the amino acid sequence designated "Comparison Protein" to the amino acid sequence designated "PRO", wherein "PRO" represents the amino acid sequence of a hypothetical PRO polypeptide of interest, "Comparison Protein" represents the amino acid sequence of a polypeptide against which the "PRO" polypeptide of interest is being compared, and "X", "Y" and "Z" each represent different hypothetical amino acid residues.

10 Unless specifically stated otherwise, all % amino acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program. However, % amino acid sequence identity values may also be obtained as described below by using the WU-BLAST-2 computer program (Altschul et al., *Methods in Enzymology* 266:460-480 (1996)). Most of the WU-BLAST-2 search parameters are set to the default values. Those not set to default values, i.e., the adjustable parameters, are set 15 with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11, and scoring matrix = BLOSUM62. When WU-BLAST-2 is employed, a % amino acid sequence identity value is determined by dividing (a) the number of matching identical amino acid residues between the amino acid sequence of the PRO polypeptide of interest having a sequence derived from the native PRO polypeptide and the comparison amino acid sequence of interest (i.e., the sequence against which the PRO polypeptide of interest is being compared which 20 may be a PRO variant polypeptide) as determined by WU-BLAST-2 by (b) the total number of amino acid residues of the PRO polypeptide of interest. For example, in the statement "a polypeptide comprising an the amino acid sequence A which has or having at least 80% amino acid sequence identity to the amino acid sequence B", the amino acid sequence A is the comparison amino acid sequence of interest and the amino acid sequence B is the amino acid sequence of the PRO polypeptide of interest.

25 Percent amino acid sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul et al., *Nucleic Acids Res.* 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from <http://www.ncbi.nlm.nih.gov> or otherwise obtained from the National Institute of Health, Bethesda, MD. NCBI-BLAST2 uses several search parameters, wherein all of those 30 search parameters are set to default values including, for example, unmask = yes, strand = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0.01, constant for multi-pass = 25, dropoff for final gapped alignment = 25 and scoring matrix = BLOSUM62.

35 In situations where NCBI-BLAST2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

100 times the fraction X/Y

where X is the number of amino acid residues scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A.

5 "PRO variant polynucleotide" or "PRO variant nucleic acid sequence" means a nucleic acid molecule which encodes an active PRO polypeptide as defined below and which has at least about 80% nucleic acid sequence identity with a nucleotide acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed
10 herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, a PRO variant polynucleotide will have at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid
15 sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95%
20 nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity with a nucleic acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or
25 without the signal sequence, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Variants do not encompass the native nucleotide sequence.

Ordinarily, PRO variant polynucleotides are at least about 30 nucleotides in length, alternatively at least about 60 nucleotides in length, alternatively at least about 90 nucleotides in length, alternatively at least about 120 nucleotides in length, alternatively at least about 150 nucleotides in length, alternatively at least about 180
30 nucleotides in length, alternatively at least about 210 nucleotides in length, alternatively at least about 240 nucleotides in length, alternatively at least about 270 nucleotides in length, alternatively at least about 300 nucleotides in length, alternatively at least about 450 nucleotides in length, alternatively at least about 600 nucleotides in length, alternatively at least about 900 nucleotides in length, or more.

35 "Percent (%) nucleic acid sequence identity" with respect to PRO-encoding nucleic acid sequences identified herein is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in the PRO nucleic acid sequence of interest, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent

nucleic acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. For purposes herein, however, % nucleic acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1 below. The ALIGN-2 sequence comparison computer program was authored by 5 Genentech, Inc. and the source code shown in Table 1 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison 10 parameters are set by the ALIGN-2 program and do not vary.

In situations where ALIGN-2 is employed for nucleic acid sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

15

$$100 \text{ times the fraction } W/Z$$

where W is the number of nucleotides scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be 20 appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C. As examples of % nucleic acid sequence identity calculations, Tables 4 and 5, demonstrate how to calculate the 25 % nucleic acid sequence identity of the nucleic acid sequence designated "Comparison DNA" to the nucleic acid sequence designated "PRO-DNA", wherein "PRO-DNA" represents a hypothetical PRO-encoding nucleic acid sequence of interest, "Comparison DNA" represents the nucleotide sequence of a nucleic acid molecule against which the "PRO-DNA" nucleic acid molecule of interest is being compared, and "N", "L" and "V" each represent different hypothetical nucleotides.

Unless specifically stated otherwise, all % nucleic acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program. However, % nucleic 30 acid sequence identity values may also be obtained as described below by using the WU-BLAST-2 computer program (Altschul et al., *Methods in Enzymology* 266:460-480 (1996)). Most of the WU-BLAST-2 search parameters are set to the default values. Those not set to default values, i.e., the adjustable parameters, are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11, and scoring matrix = BLOSUM62. When WU-BLAST-2 is employed, a % nucleic acid sequence identity value is determined 35 by dividing (a) the number of matching identical nucleotides between the nucleic acid sequence of the PRO polypeptide-encoding nucleic acid molecule of interest having a sequence derived from the native sequence PRO polypeptide-encoding nucleic acid and the comparison nucleic acid molecule of interest (i.e., the sequence against

which the PRO polypeptide-encoding nucleic acid molecule of interest is being compared which may be a variant PRO polynucleotide) as determined by WU-BLAST-2 by (b) the total number of nucleotides of the PRO polypeptide-encoding nucleic acid molecule of interest. For example, in the statement "an isolated nucleic acid molecule comprising a nucleic acid sequence A which has or having at least 80% nucleic acid sequence identity to the nucleic acid sequence B", the nucleic acid sequence A is the comparison nucleic acid molecule of interest 5 and the nucleic acid sequence B is the nucleic acid sequence of the PRO polypeptide-encoding nucleic acid molecule of interest.

Percent nucleic acid sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul et al., *Nucleic Acids Res.* 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from <http://www.ncbi.nlm.nih.gov> or otherwise obtained from the 10 National Institute of Health, Bethesda, MD. NCBI-BLAST2 uses several search parameters, wherein all of those search parameters are set to default values including, for example, unmask = yes, strand = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0.01, constant for multi-pass = 25, dropoff for final gapped alignment = 25 and scoring matrix = BLOSUM62.

In situations where NCBI-BLAST2 is employed for sequence comparisons, the % nucleic acid sequence 15 identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

$$100 \text{ times the fraction } W/Z$$

20

where W is the number of nucleotides scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C.

25

In other embodiments, PRO variant polynucleotides are nucleic acid molecules that encode an active PRO polypeptide and which are capable of hybridizing, preferably under stringent hybridization and wash conditions, to nucleotide sequences encoding a full-length PRO polypeptide as disclosed herein. PRO variant polypeptides may be those that are encoded by a PRO variant polynucleotide.

30

"Isolated," when used to describe the various polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, the polypeptide will be purified (1) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to 35 homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Isolated polypeptide includes polypeptide *in situ* within recombinant cells, since at least one component of the PRO polypeptide natural environment will not be present. Ordinarily, however, isolated

polypeptide will be prepared by at least one purification step.

An "isolated" PRO polypeptide-encoding nucleic acid or other polypeptide-encoding nucleic acid is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the polypeptide-encoding nucleic acid. An isolated polypeptide-encoding nucleic acid molecule is other than in the form or setting in which it is found in nature.

5 Isolated polypeptide-encoding nucleic acid molecules therefore are distinguished from the specific polypeptide-encoding nucleic acid molecule as it exists in natural cells. However, an isolated polypeptide-encoding nucleic acid molecule includes polypeptide-encoding nucleic acid molecules contained in cells that ordinarily express the polypeptide where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

10 The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

15 Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and 20 in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

25 The term "antibody" is used in the broadest sense and specifically covers, for example, single anti-PRO monoclonal antibodies (including agonist, antagonist, and neutralizing antibodies), anti-PRO antibody compositions with polyepitopic specificity, single chain anti-PRO antibodies, and fragments of anti-PRO antibodies (see below). The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally-occurring mutations that may be present in minor amounts.

30 "Stringency" of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured DNA to reanneal when complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature which can be used. 35 As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel et al., *Current Protocols in Molecular Biology*, Wiley Interscience Publishers, (1995).

"Stringent conditions" or "high stringency conditions", as defined herein, may be identified by those that: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50°C; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42°C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC (sodium chloride/sodium citrate) and 50% formamide at 55°C, followed by a high-stringency wash consisting of 0.1 x SSC containing EDTA at 55°C.

10 "Moderately stringent conditions" may be identified as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989, and include the use of washing solution and hybridization conditions (e.g., temperature, ionic strength and %SDS) less stringent than those described above. An example of moderately stringent conditions is overnight incubation at 37°C in a solution comprising: 20% formamide, 5 x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x 15 Denhardt's solution, 10% dextran sulfate, and 20 mg/ml denatured sheared salmon sperm DNA, followed by washing the filters in 1 x SSC at about 37-50°C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

20 The term "epitope tagged" when used herein refers to a chimeric polypeptide comprising a PRO polypeptide fused to a "tag polypeptide". The tag polypeptide has enough residues to provide an epitope against which an antibody can be made, yet is short enough such that it does not interfere with activity of the polypeptide to which it is fused. The tag polypeptide preferably also is fairly unique so that the antibody does not substantially cross-react with other epitopes. Suitable tag polypeptides generally have at least six amino acid residues and usually between about 8 and 50 amino acid residues (preferably, between about 10 and 20 amino acid residues).

25 As used herein, the term "immunoadhesin" designates antibody-like molecules which combine the binding specificity of a heterologous protein (an "adhesin") with the effector functions of immunoglobulin constant domains. Structurally, the immunoadhesins comprise a fusion of an amino acid sequence with the desired binding specificity which is other than the antigen recognition and binding site of an antibody (i.e., is "heterologous"), and an immunoglobulin constant domain sequence. The adhesin part of an immunoadhesin molecule typically is a contiguous amino acid sequence comprising at least the binding site of a receptor or a ligand. The 30 immunoglobulin constant domain sequence in the immunoadhesin may be obtained from any immunoglobulin, such as IgG-1, IgG-2, IgG-3, or IgG-4 subtypes, IgA (including IgA-1 and IgA-2), IgE, IgD or IgM.

35 "Active" or "activity" for the purposes herein refers to form(s) of a PRO polypeptide which retain a biological and/or an immunological activity of native or naturally-occurring PRO, wherein "biological" activity refers to a biological function (either inhibitory or stimulatory) caused by a native or naturally-occurring PRO other than the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring PRO and an "immunological" activity refers to the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring PRO.

The term "antagonist" is used in the broadest sense, and includes any molecule that partially or fully blocks, inhibits, or neutralizes a biological activity of a native PRO polypeptide disclosed herein. In a similar manner, the term "agonist" is used in the broadest sense and includes any molecule that mimics a biological activity of a native PRO polypeptide disclosed herein. Suitable agonist or antagonist molecules specifically include agonist or antagonist antibodies or antibody fragments, fragments or amino acid sequence variants of native PRO polypeptides, peptides, antisense oligonucleotides, small organic molecules, etc. Methods for identifying agonists or antagonists of a PRO polypeptide may comprise contacting a PRO polypeptide with a candidate agonist or antagonist molecule and measuring a detectable change in one or more biological activities normally associated with the PRO polypeptide.

"Treatment" refers to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or slow down (lessen) the targeted pathologic condition or disorder. Those in need of treatment include those already with the disorder as well as those prone to have the disorder or those in whom the disorder is to be prevented.

"Chronic" administration refers to administration of the agent(s) in a continuous mode as opposed to an acute mode, so as to maintain the initial therapeutic effect (activity) for an extended period of time. "Intermittent" administration is treatment that is not consecutively done without interruption, but rather is cyclic in nature.

"Mammal" for purposes of treatment refers to any animal classified as a mammal, including humans, domestic and farm animals, and zoo, sports, or pet animals, such as dogs, cats, cattle, horses, sheep, pigs, goats, rabbits, etc. Preferably, the mammal is human.

Administration "in combination with" one or more further therapeutic agents includes simultaneous (concurrent) and consecutive administration in any order.

"Carriers" as used herein include pharmaceutically acceptable carriers, excipients, or stabilizers which are nontoxic to the cell or mammal being exposed thereto at the dosages and concentrations employed. Often the physiologically acceptable carrier is an aqueous pH buffered solution. Examples of physiologically acceptable carriers include buffers such as phosphate, citrate, and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptide; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN™, polyethylene glycol (PEG), and PLURONICS™.

"Antibody fragments" comprise a portion of an intact antibody, preferably the antigen binding or variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', F(ab')₂, and Fv fragments; diabodies; linear antibodies (Zapata et al., *Protein Eng.* 8(10): 1057-1062 [1995]); single-chain antibody molecules; and multispecific antibodies formed from antibody fragments.

Papain digestion of antibodies produces two identical antigen-binding fragments, called "Fab" fragments, each with a single antigen-binding site, and a residual "Fc" fragment, a designation reflecting the ability to crystallize readily. Pepsin treatment yields an F(ab')₂ fragment that has two antigen-combining sites and is still

capable of cross-linking antigen.

"Fv" is the minimum antibody fragment which contains a complete antigen-recognition and -binding site. This region consists of a dimer of one heavy- and one light-chain variable domain in tight, non-covalent association. It is in this configuration that the three CDRs of each variable domain interact to define an antigen-binding site on the surface of the V_H - V_L dimer. Collectively, the six CDRs confer antigen-binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen) has the ability to recognize and bind antigen, although at a lower affinity than the entire binding site.

The Fab fragment also contains the constant domain of the light chain and the first constant domain (CH1) of the heavy chain. Fab fragments differ from Fab' fragments by the addition of a few residues at the carboxy terminus of the heavy chain CH1 domain including one or more cysteines from the antibody hinge region. 10 Fab'-SH is the designation herein for Fab' in which the cysteine residue(s) of the constant domains bear a free thiol group. $F(ab')_2$ antibody fragments originally were produced as pairs of Fab' fragments which have hinge cysteines between them. Other chemical couplings of antibody fragments are also known.

The "light chains" of antibodies (immunoglobulins) from any vertebrate species can be assigned to one of two clearly distinct types, called kappa and lambda, based on the amino acid sequences of their constant 15 domains.

Depending on the amino acid sequence of the constant domain of their heavy chains, immunoglobulins can be assigned to different classes. There are five major classes of immunoglobulins: IgA, IgD, IgE, IgG, and IgM, and several of these may be further divided into subclasses (isotypes), e.g., IgG1, IgG2, IgG3, IgG4, IgA, and IgA2.

20 "Single-chain Fv" or "sFv" antibody fragments comprise the V_H and V_L domains of antibody, wherein these domains are present in a single polypeptide chain. Preferably, the Fv polypeptide further comprises a polypeptide linker between the V_H and V_L domains which enables the sFv to form the desired structure for antigen binding. For a review of sFv, see Pluckthun in The Pharmacology of Monoclonal Antibodies, vol. 113, Rosenberg and Moore eds., Springer-Verlag, New York, pp. 269-315 (1994).

25 The term "diabodies" refers to small antibody fragments with two antigen-binding sites, which fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) in the same polypeptide chain (V_H - V_L). By using a linker that is too short to allow pairing between the two domains on the same chain, the domains are forced to pair with the complementary domains of another chain and create two antigen-binding sites. Diabodies are described more fully in, for example, EP 404,097; WO 93/11161; and 30 Hollinger et al., Proc. Natl. Acad. Sci. USA, 90:6444-6448 (1993).

An "isolated" antibody is one which has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials which would interfere with diagnostic or therapeutic uses for the antibody, and may include enzymes, hormones, and other proteinaceous or nonproteinaceous solutes. In preferred embodiments, the antibody will be purified (1) to greater than 95% by weight of antibody as determined by the Lowry method, and most preferably more than 99% by weight, (2) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (3) to homogeneity by SDS-PAGE under reducing or nonreducing

conditions using Coomassie blue or, preferably, silver stain. Isolated antibody includes the antibody *in situ* within recombinant cells since at least one component of the antibody's natural environment will not be present. Ordinarily, however, isolated antibody will be prepared by at least one purification step.

An antibody that "specifically binds to" or is "specific for" a particular polypeptide or an epitope on a particular polypeptide is one that binds to that particular polypeptide or epitope on a particular polypeptide without substantially binding to any other polypeptide or polypeptide epitope.

The word "label" when used herein refers to a detectable compound or composition which is conjugated directly or indirectly to the antibody so as to generate a "labeled" antibody. The label may be detectable by itself (e.g. radioisotope labels or fluorescent labels) or, in the case of an enzymatic label, may catalyze chemical alteration of a substrate compound or composition which is detectable.

By "solid phase" is meant a non-aqueous matrix to which the antibody of the present invention can adhere. Examples of solid phases encompassed herein include those formed partially or entirely of glass (e.g., controlled pore glass), polysaccharides (e.g., agarose), polyacrylamides, polystyrene, polyvinyl alcohol and silicones. In certain embodiments, depending on the context, the solid phase can comprise the well of an assay plate; in others it is a purification column (e.g., an affinity chromatography column). This term also includes a discontinuous solid phase of discrete particles, such as those described in U.S. Patent No. 4,275,149.

A "liposome" is a small vesicle composed of various types of lipids, phospholipids and/or surfactant which is useful for delivery of a drug (such as a PRO polypeptide or antibody thereto) to a mammal. The components of the liposome are commonly arranged in a bilayer formation, similar to the lipid arrangement of biological membranes.

A "small molecule" is defined herein to have a molecular weight below about 500 Daltons.

An "effective amount" of a polypeptide disclosed herein or an agonist or antagonist thereof is an amount sufficient to carry out a specifically stated purpose. An "effective amount" may be determined empirically and in a routine manner, in relation to the stated purpose.

Table 1

Table 1 (cont')

```

/*
 */
#include <stdio.h>
#include <ctype.h>

5   #define MAXJMP      16    /* max jumps in a diag */
#define MAXGAP      24    /* don't continue to penalize gaps larger than this */
#define J MPS      1024   /* max jmps in an path */
#define MX       4     /* save if there's at least MX-1 bases since last jmp */

10  #define DMAT      3     /* value of matching bases */
#define DMIS      0     /* penalty for mismatched bases */
#define DINS0     8     /* penalty for a gap */
#define DINS1     1     /* penalty per base */
15  #define PINS0     8     /* penalty for a gap */
#define PINS1     4     /* penalty per residue */

20  struct jmp {
        short      n[MAXJMP];    /* size of jmp (neg for delay) */
        unsigned short x[MAXJMP]; /* base no. of jmp in seq x */
        };                      /* limits seq to 2^16 - 1 */

25  struct diag {
        int       score;        /* score at last jmp */
        long      offset;       /* offset of prev block */
        short      ijmp;        /* current jmp index */
        struct jmp  jp;         /* list of jmps */
        };

30  struct path {
        int       spc;          /* number of leading spaces */
        short      n[J MPS];    /* size of jmp (gap) */
        int       x[J MPS];    /* loc of jmp (last elem before gap) */
        };

35  char      *ofile;        /* output file name */
char      *namex[2];    /* seq names: getseqs() */
char      *prog;          /* prog name for err msgs */
char      *seqx[2];        /* seqs: getseqs() */
40  int       dmax;          /* best diag: nw() */
int       dmax0;         /* final diag */
int       dna;            /* set if dna: main() */
int       endgaps;        /* set if penalizing end gaps */
int       gapx, gapy;    /* total gaps in seqs */
45  int       len0, len1;    /* seq lens */
int       ngapx, ngapy;  /* total size of gaps */
int       smax;           /* max score: nw() */
int       *xbm;            /* bitmap for matching */
long      offset;         /* current offset in jmp file */
50  struct diag  *dx;        /* holds diagonals */
struct path   pp[2];    /* holds path for seqs */

55  char      *calloc(), *malloc(), *index(), *strcpy();
char      *getseq(), *g_malloc();

```

Table 1 (cont')

```

/* Needleman-Wunsch alignment program
*
* usage: progs file1 file2
* where file1 and file2 are two dna or two protein sequences.
* The sequences can be in upper- or lower-case and may contain ambiguity
* Any lines beginning with ';', '>' or '<' are ignored
* Max file length is 65535 (limited by unsigned short x in the jmp struct)
* A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
* Output is in the file "align.out"
*
5
* The program may create a tmp file in /tmp to hold info about traceback.
* Original version developed under BSD 4.3 on a vax 8650
*/
10
#include "nw.h"
15
#include "day.h"

20
static _dbval[26] = {
    1,14,2,13,0,0,4,11,0,0,12,0,3,15,0,0,0,5,6,8,8,7,9,0,10,0
};

25
static _pbval[26] = {
    1, 2|(1<<('D'-'A'))|(1<<('N'-'A')), 4, 8, 16, 32, 64,
    128, 256, 0xFFFFFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
    1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
    1<<23, 1<<24, 1<<25|(1<<('E'-'A'))|(1<<('Q'-'A'))
};

30
main(ac, av)
    int ac;
    char *av[];
{
    prog = av[0];
    if (ac != 3) {
        35
        fprintf(stderr, "usage: %s file1 file2\n", prog);
        fprintf(stderr, "where file1 and file2 are two dna or two protein sequences.\n");
        fprintf(stderr, "The sequences can be in upper- or lower-case\n");
        fprintf(stderr, "Any lines beginning with ';' or '<' are ignored\n");
        fprintf(stderr, "Output is in the file \"align.out\"\n");
        exit(1);
    }
    40
    namex[0] = av[1];
    namex[1] = av[2];
    seqx[0] = getseq(namex[0], &len0);
    seqx[1] = getseq(namex[1], &len1);
    xbm = (dna)? _dbval : _pbval;

    endgaps = 0; /* 1 to penalize endgaps */
    ofile = "align.out"; /* output file */

50
    nw(); /* fill in the matrix, get the possible jmps */
    readjmps(); /* get the actual jmps */
    print(); /* print stats, alignment */

    55
    cleanup(0); /* unlink any tmp files */
}

```

Table 1 (cont')

```

/* do the alignment, return best score: main()
 * dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
 * pro: PAM 250 values
 * When scores are equal, we prefer mismatches to any gap, prefer
 5  * a new gap to extending an ongoing gap, and prefer a gap in seqx
 * to a gap in seq y.
 */
nw()
{
10    char *          *px, *py;      /* seqs and ptrs */
    int      *ndely, *dely;      /* keep track of dely */
    int      ndelx, delx;      /* keep track of delx */
    int      *tmp;            /* for swapping row0, row1 */
    int      mis;             /* score for each type */
    int      ins0, ins1;      /* insertion penalties */
15    register id;           /* diagonal index */
    register ij;            /* jmp index */
    register *col0, *col1;    /* score for curr, last row */
    register xx, yy;         /* index into seqs */
20
    dx = (struct diag *)g_calloc("to get diags", len0+len1+1, sizeof(struct diag));
    ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
    dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
25    col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
    col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
    ins0 = (dna)? DINS0 : PINS0;
    ins1 = (dna)? DINS1 : PINS1;
30
    smax = -10000;
    if (endgaps) {
        for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
            col0[yy] = dely[yy] = col0[yy-1] - ins1;
            ndely[yy] = yy;
        }
        col0[0] = 0;      /* Waterman Bull Math Biol 84 */
    }
35    else
        for (yy = 1; yy <= len1; yy++)
            dely[yy] = -ins0;
40
        /* fill in match matrix
        */
45    for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
        /* initialize first entry in col
        */
        if (endgaps) {
            if (xx == 1)
                col1[0] = delx = -(ins0+ins1);
            else
                col1[0] = delx = col0[0] - ins1;
                ndelx = xx;
        }
50        else {
            col1[0] = 0;
            delx = -ins0;
            ndelx = 0;
        }
55
        }
60
    }

```

Table 1 (cont')

```

...nw

for (py = seqx[1], yy = 1; yy <= len1; py++, yy++) {
    mis = col0[yy-1];
    if (dna)
        mis += (xbm[*px-'A']&xbm[*py-'A'])? DMAT : DMIS;
    else
        mis += _day[*px-'A'][*py-'A'];

    /* update penalty for del in x seq;
     * favor new del over ongoing del
     * ignore MAXGAP if weighting endgaps
     */
    if (endgaps || ndely[yy] < MAXGAP) {
        if (col0[yy] - ins0 >= dely[yy]) {
            dely[yy] = col0[yy] - (ins0+ins1);
            ndely[yy] = 1;
        } else {
            dely[yy] -= ins1;
            ndely[yy]++;
        }
    } else {
        if (col0[yy] - (ins0+ins1) >= dely[yy]) {
            dely[yy] = col0[yy] - (ins0+ins1);
            ndely[yy] = 1;
        } else
            ndely[yy]++;
    }

    /* update penalty for del in y seq;
     * favor new del over ongoing del
     */
    if (endgaps || ndelx < MAXGAP) {
        if (col1[yy-1] - ins0 >= delx) {
            delx = col1[yy-1] - (ins0+ins1);
            ndelx = 1;
        } else {
            delx -= ins1;
            ndelx++;
        }
    } else {
        if (col1[yy-1] - (ins0+ins1) >= delx) {
            delx = col1[yy-1] - (ins0+ins1);
            ndelx = 1;
        } else
            ndelx++;
    }

    /* pick the maximum score; we're favoring
     * mis over any del and delx over dely
     */
}

```

55

60

Table 1 (cont')

```

...nw

5      id = xx - yy + len1 - 1;
      if (mis >= delx && mis >= dely[yy])
          col1[yy] = mis;
      else if (delx >= dely[yy]) {
          col1[yy] = delx;
          ij = dx[id].ijmp;
          if (dx[id].jp.n[0] && (!dna || (ndelx >= MAXJMP
10      && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
              dx[id].ijmp++;
              if (++ij >= MAXJMP) {
                  writejmps(id);
                  ij = dx[id].ijmp = 0;
                  dx[id].offset = offset;
                  offset += sizeof(struct jmp) + sizeof(offset);
15      }
              }
          dx[id].jp.n[ij] = ndelx;
          dx[id].jp.x[ij] = xx;
          dx[id].score = delx;
      }
      else {
          col1[yy] = dely[yy];
          ij = dx[id].ijmp;
20      if (dx[id].jp.n[0] && (!dna || (ndely[yy] >= MAXJMP
          && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
              dx[id].ijmp++;
              if (++ij >= MAXJMP) {
                  writejmps(id);
                  ij = dx[id].ijmp = 0;
                  dx[id].offset = offset;
                  offset += sizeof(struct jmp) + sizeof(offset);
30      }
              }
          dx[id].jp.n[ij] = -ndely[yy];
          dx[id].jp.x[ij] = xx;
          dx[id].score = dely[yy];
      }
      if (xx == len0 && yy < len1) {
40      /* last col
          */
          if (endgaps)
              col1[yy] -= ins0+ins1*(len1-yy);
          if (col1[yy] > smax) {
45          smax = col1[yy];
              dmax = id;
          }
      }
      if (endgaps && xx < len0)
          col1[yy-1] -= ins0+ins1*(len0-xx);
      if (col1[yy-1] > smax) {
          smax = col1[yy-1];
          dmax = id;
55      }
      tmp = col0; col0 = col1; col1 = tmp;
      }
      (void) free((char *)ndely);
      (void) free((char *)dely);
      (void) free((char *)col0);
60      (void) free((char *)col1);
  }

```

Table 1 (cont')

```

/*
*
* print() -- only routine visible outside this module
*
5  * static:
* getmat() -- trace back best path, count matches: print()
* pr_align() -- print alignment of described in array p[]: print()
* dumpblock() -- dump a block of lines with numbers, stars: pr_align()
* nums() -- put out a number line: dumpblock()
10 * putline() -- put out a line (name, [num], seq, [num]): dumpblock()
* stars() - put a line of stars: dumpblock()
* stripname() -- strip any path and prefix from a seqname
*/
15 #include "nw.h"

#define SPC      3
#define P_LINE   256 /* maximum output line */
#define P_SPC    3 /* space between name or num and seq */
20
extern _day[26][26];
int olen;          /* set output line length */
FILE *fx;          /* output file */

25 print()
{
    int lx, ly, firstgap, lastgap; /* overlap */

    if ((fx = fopen(ofile, "w")) == 0) {
30        fprintf(stderr, "%s: can't write %s\n", prog, ofile);
        cleanup(1);
    }
    fprintf(fx, "<first sequence: %s (length = %d)\n", namex[0], len0);
    fprintf(fx, "<second sequence: %s (length = %d)\n", namex[1], len1);
    olen = 60;
    lx = len0;
    ly = len1;
    firstgap = lastgap = 0;
    if (dmax < len1 - 1) { /* leading gap in x */
40        pp[0].spc = firstgap = len1 - dmax - 1;
        ly -= pp[0].spc;
    }
    else if (dmax > len1 - 1) { /* leading gap in y */
        pp[1].spc = firstgap = dmax - (len1 - 1);
        lx -= pp[1].spc;
    }
    if (dmax0 < len0 - 1) { /* trailing gap in x */
50        lastgap = len0 - dmax0 - 1;
        lx -= lastgap;
    }
    else if (dmax0 > len0 - 1) { /* trailing gap in y */
        lastgap = dmax0 - (len0 - 1);
        ly -= lastgap;
    }
55    getmat(lx, ly, firstgap, lastgap);
    pr_align();
}

```

Table 1 (cont?)

```

/*
 * trace back the best path, count matches
 */
static
5  getmat(lx, ly, firstgap, lastgap)           getmat
    int      lx, ly;                      /* "core" (minus endgaps) */
    int      firstgap, lastgap;           /* leading/trailing overlap */
{
    int      nm, i0, i1, siz0, siz1;
10   char     outx[32];
    double    pct;
    register  n0, n1;
    register char *p0, *p1;

15   /* get total matches, score
 */
    i0 = i1 = siz0 = siz1 = 0;
    p0 = seqx[0] + pp[1].spc;
    p1 = seqx[1] + pp[0].spc;
20   n0 = pp[1].spc + 1;
    n1 = pp[0].spc + 1;

    nm = 0;
25   while ( *p0 && *p1 ) {
        if (siz0) {
            p1++;
            n1++;
            siz0--;
        }
30   else if (siz1) {
            p0++;
            n0++;
            siz1--;
        }
35   else {
            if (xbm[*p0-'A']&xbm[*p1-'A'])
                nm++;
            if (n0++ == pp[0].n[i0])
                siz0 = pp[0].n[i0++];
            if (n1++ == pp[1].n[i1])
                siz1 = pp[1].n[i1++];
            p0++;
            p1++;
        }
45   }

    /* pct homology:
     * if penalizing endgaps, base is the shorter seq
     * else, knock off overhangs and take shorter core
     */
50   if (endgaps)
        lx = (len0 < len1)? len0 : len1;
    else
        lx = (lx < ly)? lx : ly;
55   pct = 100.* (double)nm/(double)lx;
    fprintf(fx, "\n");
    fprintf(fx, "<%d match%s in an overlap of %d: %.2f percent similarity\n",
            nm, (nm == 1)? "" : "es", lx, pct);

```

60

Table 1 (cont')

```

5   fprintf(fx, "<gaps in first sequence: %d", gapx); ...getmat
  if (gapx) {
    (void) sprintf(outx, " (%d %s%s)",
      ngapx, (dna)? "base": "residue", (ngapx == 1)? ":" "s");
    sprintf(fx, "%s", outx);

10  fprintf(fx, ", gaps in second sequence: %d", gapy);
  if (gapy) {
    (void) sprintf(outx, " (%d %s%s",
      ngapy, (dna)? "base": "residue", (ngapy == 1)? ":" "s");
    sprintf(fx, "%s", outx);
  }
  if (dna)
    fprintf(fx,
      "\n<score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
      smax, DMAT, DMIS, DINS0, DINS1);
  else
    fprintf(fx,
      "\n<score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n",
      smax, PINS0, PINS1);
20  if (endgaps)
    fprintf(fx,
      "<endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s\n",
      firstgap, (dna)? "base": "residue", (firstgap == 1)? ":" "s",
      lastgap, (dna)? "base": "residue", (lastgap == 1)? ":" "s");
  else
    fprintf(fx, "<endgaps not penalized\n");
}

30  static nm;          /* matches in core -- for checking */
  static lmax;         /* lengths of stripped file names */
  static ij[2];        /* jmp index for a path */
  static nc[2];        /* number at start of current line */
  static ni[2];        /* current elem number -- for gapping */
35  static siz[2];
  static char *ps[2];  /* ptr to current element */
  static char *po[2];  /* ptr to next output char slot */
  static char out[2][P_LINE]; /* output line */
  static char star[P_LINE]; /* set by stars() */
40
/*
 * print alignment of described in struct path pp[]
 */
static
45  pr_align0
{
  int nn;          /* char count */
  int more;
  register i;
50
  for (i = 0, lmax = 0; i < 2; i++) {
    nn = stripname(namef[i]);
    if (nn > lmax)
      lmax = nn;
55
    nc[i] = 1;
    ni[i] = 1;
    siz[i] = ij[i] = 0;
    ps[i] = seqx[i];
    po[i] = out[i];
}

```

Table 1 (cont')

```

for (nn = nm = 0, more = 1; more; ) {
    for (i = more - 0; i < 2; i++) {
        /*
         * do we have more of this sequence?
         */
        if (!*ps[i])
            continue;
    more++;
    if (pp[i].spc) { /* leading space */
        *po[i]++ = ' ';
        pp[i].spc--;
    }
    else if (siz[i]) { /* in a gap */
        *po[i]++ = ' ';
        siz[i]--;
    }
    else { /* we're putting a seq element
        */
        *po[i] = *ps[i];
        if (islower(*ps[i]))
            *ps[i] = toupper(*ps[i]);
        po[i]++;
        ps[i]++;
        /*
         * are we at next gap for this seq?
         */
        if (ni[i] == pp[i].x[ij[i]]) {
            /*
             * we need to merge all gaps
             * at this location
             */
            siz[i] = pp[i].n[ij[i]]++;
            while (ni[i] == pp[i].x[ij[i]])
                siz[i] += pp[i].n[ij[i]]++;
        }
        ni[i]++;
    }
    if (++nn == olen || !more && nm) {
        dumpblock();
        for (i = 0; i < 2; i++)
            po[i] = out[i];
        nn = 0;
    }
}
/*
 * dump a block of lines, including numbers, stars: pr_align()
 */
55 static
dumpblock()
{
    register i;
    for (i = 0; i < 2; i++)
        *po[i]-- = '\0';
    ...
}

```

dumpblock

Table 1 (cont?)

```

...dumpblock

5   (void) putc('u', fx);
    for (i = 0; i < 2; i++) {
        if (*out[i] && (*out[i] != ' ' || *(po[i]) != ' ')) {
            if (i == 0)
                nums(i);
            if (i == 0 && *out[1])
                stars();
10   putline(ix);
            if (i == 0 && *out[1])
                fprintf(fx, star);
            if (i == 1)
                nums(i);
15   }
    }

20   /*
21   * put out a number line: dumpblock()
22   */
23   static
24   nums(ix)
25   {
        int      ix;      /* index in out[] holding seq line */
        num
26   char      nline[P_LINE];
27   register  i, j;
28   register char  *pn, *px, *py;

30   for (pn = nline, i = 0; i < lmax+P_SPC; i++, pn++)
        *pn = ' ';
29   for (i = nc[ix], py = out[ix]; *py; py++, pn++) {
        if (*py == ' ' || *py == '-')
            *pn = ' ';
35   else {
            if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                j = (i < 0)? -i : i;
                for (px = pn; j /= 10, px--)
                    *px = j%10 + '0';
36   if (i < 0)
                    *px = '-';
37   }
            else
                *pn = ' ';
38   i++;
39   }
40   }
41   *pn = '\0';
42   nc[ix] = i;
43   for (pn = nline; *pn; pn++)
44     (void) putc(*pn, fx);
45   (void) putc('\n', fx);
    }

55   /*
56   * put out a line (name, [num], seq, [num]): dumpblock()
57   */
58   static
59   putline(ix)
60   int      ix;
    {

```

Table 1 (cont')

```

...putline
5
int          i;
register char *px;
for (px = namex[ix], i = 0; *px && *px != ':'; px++, i++)
    (void) putc(*px, fx);
for (; i < lmax+P_SPC; i++)
    (void) putc(' ', fx);
10
/* these count from 1:
 * ni[] is current element (from 1)
 * nc[] is number at start of current line
 */
15
for (px = out[ix]; *px; px++)
    (void) putc(*px&0x7F, fx);
    (void) putc('\n', fx);
}
20
/*
 * put a line of stars (seqs always in out[0], out[1]): dumpblock()
 */
static
25 stars()
{
    int          i;
    register char *p0, *p1, cx, *px;
30
if (!*out[0] || (*out[0] == ' ' && *(po[0]) == ' ') ||
    !*out[1] || (*out[1] == ' ' && *(po[1]) == ' '))
    return;
px = star;
for (i = lmax+P_SPC; i; i--)
    *px++ = ' ';
35
for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
    if (isalpha(*p0) && isalpha(*p1)) {
40
        if (xbm[*p0-'A']&xbm[*p1-'A']) {
            cx = '*';
            nm++;
        }
        else if (!dma && _day[*p0-'A'][*p1-'A'] > 0)
            cx = '.';
        else
            cx = ' ';
    }
    else
        cx = ' ';
50
    *px++ = cx;
    *px++ = cx;
}
*px++ = '\n';
*px = '\0';
55 }

```

Table 1 (cont')

```

/*
 * strip path or prefix from pn, return len: pr_align()
 */
static
5  stripname(pn)
    char    *pn;      /* file name (may be path) */
{
    register char    *px, *py;

10   py = 0;
    for (px = pn; *px; px++)
        if (*px == '/')
            py = px + 1;
    if (py)
15     (void) strcpy(pn, py);
    return(strlen(pn));
}
20

```

Table 1 (cont')

```

/*
 * cleanup() -- cleanup any tmp file
 * getseq() -- read in seq, set dna, len, maxlen
 * g_calloc() -- calloc() with error checkin
 5   * readjmps() -- get the good jmps, from tmp file if necessary
 * writejmps() -- write a filled array of jmps to a tmp file: nw()
 */
#include "nw.h"
#include <sys/file.h>
10
char  *jname = "/tmp/homgXXXXXX";           /* tmp file for jmps */
FILE  *fj;
15
int   cleanup();                         /* cleanup tmp file */
long  lseek();
1
/*
 * remove any tmp file if we blow
 */
20
cleanup(i)
    int   i;
{
    if (fj)
        (void) unlink(jname);
25
    exit(i);
}

/*
 * read, return ptr to seq, set dna, len, maxlen
30   * skip lines starting with ';', '<', or '>'
 * seq in upper or lower case
 */
char  *
getseq(file, len)
35
    char  *file;    /* file name */
    int   *len;     /* seq len */
{
    char  line[1024], *pseq;
    register char  *px, *py;
40
    int   natgc, tlen;
    FILE  *fp;
45
    if ((fp = fopen(file, "r")) == 0) {
        fprintf(stderr, "%s: can't read %s\n", prog, file);
        exit(1);
    }
    tlen = natgc = 0;
    while (fgets(line, 1024, fp)) {
50
        if (*line == ';' || *line == '<' || *line == '>')
            continue;
        for (px = line; *px != '\n'; px++)
            if (isupper(*px) || islower(*px))
                tlen++;
    }
55
    if ((pseq = malloc((unsigned)(tlen+6))) == 0) {
        fprintf(stderr, "%s: malloc() failed to get %d bytes for %s\n", prog, tlen+6, file);
        exit(1);
    }
    pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';
60

```

Table 1 (cont')

```

...getseq

5      py = pseq + 4;
*len = tlen;
rewind(fp);

10     while (fgets(line, 1024, fp)) {
        if (*line == ';' || *line == '<' || *line == '>')
            continue;
        for (px = line; *px != '\n'; px++) {
            if (isupper(*px))
                *py++ = *px;
            else if (islower(*px))
                *py++ = toupper(*px);
            if (index("ATGCU", *(py-1)))
                natgc++;
        }
        *py++ = '\0';
*py = '\0';
20     (void) fclose(fp);
dna = natgc > (tlen/3);
return(pseq+4);
}

25     char * g_malloc(msg, nx, sz)
30     {
        char *msg; /* program, calling routine */
        int nx, sz; /* number and size of elements */
        char *px, *calloc();
        if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
            if (*msg) {
                fprintf(stderr, "%s: g_malloc() failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);
                exit(1);
            }
        }
        return(px);
    }
40     /*
        * get final jmps from dx[] or tmp file, set pp[], reset dmax: main()
        */
45     readjmps()
{
        int fd = -1;
        int siz, i0, i1;
        register i, j, xx;
50     if (fj) {
            (void) fclose(fj);
            if ((fd = open(jname, O_RDONLY, 0)) < 0) {
                fprintf(stderr, "%s: can't open() %s\n", prog, jname);
                cleanup(1);
            }
        }
        for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; ; i++) {
            while (1)
                for (j = dx[dmax].jmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)
60
;
```

Table 1 (cont')

```

...readjmps
5
if (j < 0 && dx[dmax].offset && fj) {
    (void) lseek(fd, dx[dmax].offset, 0);
    (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp));
    (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset));
    dx[dmax].jmp = MAXJMP-1;
}
else
    break;
10
}
if (i >= JMPS) {
    fprintf(stderr, "%s: too many gaps in alignment\n", prog);
    cleanup(1);
}
15
if (j >= 0) {
    siz = dx[dmax].jp.n[j];
    xx = dx[dmax].jp.x[j];
    dmax += siz;
    if (siz < 0) { /* gap in second seq */
        pp[1].n[i1] = -siz;
        xx += siz;
        /* id = xx - yy + len1 - 1
        */
        pp[1].x[i1] = xx - dmax + len1 - 1;
        gapy++;
        ngapy -= siz;
    }
    /* ignore MAXGAP when doing endgaps */
    siz = (-siz < MAXGAP || endgaps)? -siz : MAXGAP;
    i1++;
}
20
}
else if (siz > 0) { /* gap in first seq */
    pp[0].n[i0] = siz;
    pp[0].x[i0] = xx;
    gapx++;
    ngapx += siz;
}
25
/* ignore MAXGAP when doing endgaps */
siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
i0++;
}
30
}
else if (siz > 0) { /* gap in first seq */
    pp[0].n[i0] = siz;
    pp[0].x[i0] = xx;
    gapx++;
    ngapx += siz;
}
35
/* ignore MAXGAP when doing endgaps */
siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
i0++;
}
40
}
else
    break;
}
45
/* reverse the order of jmps
 */
for (j = 0, i0--; j < i0; j++, i0--) {
    i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
    i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
}
50
for (j = 0, i1--; j < i1; j++, i1--) {
    i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
    i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
}
55
if (fd >= 0)
    (void) close(fd);
if (fj) {
    (void) unlink(jname);
    fj = 0;
    offset = 0;
}
60
}
}

```

Table 1 (cont')

```

/*
 * write a filled jmp struct offset of the prev one (if any): nw0
 */
5   writejmps(ix)                                writejmps
      int      ix;
{
      char    *mktemp();
10  if (!fj) {
          if (mktemp(jname) < 0) {
              fprintf(stderr, "%s: can't mktemp() %s\n", prog, jname);
              cleanup(1);
          }
15  if ((fj = fopen(jname, "w")) == 0) {
              fprintf(stderr, "%s: can't write %s\n", prog, jname);
              exit(1);
          }
20  (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);
  (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);
}

25

30

35

40

45

50

55

60

```

Table 2

PRO	XXXXXXXXXXXXXXXXXXXX	(Length = 15 amino acids)
Comparison Protein	XXXXXXYYYYYYYY	(Length = 12 amino acids)

5 % amino acid sequence identity =

(the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) =

10 5 divided by 15 = 33.3%

Table 3

PRO	XXXXXXXXXXXX	(Length = 10 amino acids)
Comparison Protein	XXXXXXYYYYYYZZYZ	(Length = 15 amino acids)

% amino acid sequence identity =

(the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) =

20 5 divided by 10 = 50%

Table 4

25	PRO-DNA	NNNNNNNNNNNNNN	(Length = 14 nucleotides)
	Comparison DNA	NNNNNNLLLLLLLL	(Length = 16 nucleotides)

30 % nucleic acid sequence identity =

(the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) =

35 6 divided by 14 = 42.9%

Table 5

PRO-DNA	NNNNNNNNNNNN	(Length = 12 nucleotides)
Comparison DNA	NNNNLLLVV	(Length = 9 nucleotides)

5 % nucleic acid sequence identity =

(the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) =

10 4 divided by 12 = 33.3%

II. Compositions and Methods of the Invention

A. Full-Length PRO Polypeptides

15 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO polypeptides. In particular, cDNAs encoding various PRO polypeptides have been identified and isolated, as disclosed in further detail in the Examples below. It is noted that proteins produced in separate expression rounds may be given different PRO numbers but the UNQ number is unique for any given DNA and the encoded protein, and will not be changed. However, for sake of simplicity, in the present specification the protein encoded by the full length native nucleic acid molecules disclosed herein 20 as well as all further native homologues and variants included in the foregoing definition of PRO, will be referred to as "PRO/number", regardless of their origin or mode of preparation.

25 As disclosed in the Examples below, various cDNA clones have been deposited with the ATCC. The actual nucleotide sequences of those clones can readily be determined by the skilled artisan by sequencing of the deposited clone using routine methods in the art. The predicted amino acid sequence can be determined from the nucleotide sequence using routine skill. For the PRO polypeptides and encoding nucleic acids described herein, Applicants have identified what is believed to be the reading frame best identifiable with the sequence information available at the time.

B. PRO Polypeptide Variants

30 In addition to the full-length native sequence PRO polypeptides described herein, it is contemplated that PRO variants can be prepared. PRO variants can be prepared by introducing appropriate nucleotide changes into the PRO DNA, and/or by synthesis of the desired PRO polypeptide. Those skilled in the art will appreciate that amino acid changes may alter post-translational processes of the PRO, such as changing the number or position of glycosylation sites or altering the membrane anchoring characteristics.

35 Variations in the native full-length sequence PRO or in various domains of the PRO described herein, can be made, for example, using any of the techniques and guidelines for conservative and non-conservative

mutations set forth, for instance, in U.S. Patent No. 5,364,934. Variations may be a substitution, deletion or insertion of one or more codons encoding the PRO that results in a change in the amino acid sequence of the PRO as compared with the native sequence PRO. Optionally the variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the PRO. Guidance in determining which amino acid residue may be inserted, substituted or deleted without adversely affecting the desired activity may be found by 5 comparing the sequence of the PRO with that of homologous known protein molecules and minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Insertions or deletions may optionally be in the range of about 1 to 5 amino acids. The variation allowed may be determined by 10 systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity exhibited by the full-length or mature native sequence.

PRO polypeptide fragments are provided herein. Such fragments may be truncated at the N-terminus or C-terminus, or may lack internal residues, for example, when compared with a full length native protein. Certain fragments lack amino acid residues that are not essential for a desired biological activity of the PRO 15 polypeptide.

PRO fragments may be prepared by any of a number of conventional techniques. Desired peptide fragments may be chemically synthesized. An alternative approach involves generating PRO fragments by enzymatic digestion, e.g., by treating the protein with an enzyme known to cleave proteins at sites defined by particular amino acid residues, or by digesting the DNA with suitable restriction enzymes and isolating the desired 20 fragment. Yet another suitable technique involves isolating and amplifying a DNA fragment encoding a desired polypeptide fragment, by polymerase chain reaction (PCR). Oligonucleotides that define the desired termini of the DNA fragment are employed at the 5' and 3' primers in the PCR. Preferably, PRO polypeptide fragments share at least one biological and/or immunological activity with the native PRO polypeptide disclosed herein.

In particular embodiments, conservative substitutions of interest are shown in Table 6 under the heading 25 of preferred substitutions. If such substitutions result in a change in biological activity, then more substantial changes, denominated exemplary substitutions in Table 6, or as further described below in reference to amino acid classes, are introduced and the products screened.

Table 6

	<u>Original Residue</u>	<u>Exemplary Substitutions</u>	<u>Preferred Substitutions</u>
5	Ala (A)	val; leu; ile	val
	Arg (R)	lys; gln; asn	lys
	Asn (N)	gln; his; lys; arg	gln
	Asp (D)	glu	glu
	Cys (C)	ser	ser
10	Gln (Q)	asn	asn
	Glu (E)	asp	asp
	Gly (G)	pro; ala	ala
	His (H)	asn; gln; lys; arg	arg
	Ile (I)	leu; val; met; ala; phe; norleucine	leu
15	Leu (L)	norleucine; ile; val; met; ala; phe	ile
	Lys (K)	arg; gln; asn	arg
	Met (M)	leu; phe; ile	leu
20	Phe (F)	leu; val; ile; ala; tyr	leu
	Pro (P)	ala	ala
	Ser (S)	thr	thr
	Thr (T)	ser	ser
	Trp (W)	tyr; phe	tyr
25	Tyr (Y)	trp; phe; thr; ser	phe
	Val (V)	ile; leu; met; phe; ala; norleucine	leu

30 Substantial modifications in function or immunological identity of the PRO polypeptide are accomplished by selecting substitutions that differ significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Naturally occurring residues are divided into groups based on common side-chain properties:

35 (1) hydrophobic: norleucine, met, ala, val, leu, ile;
 (2) neutral hydrophilic: cys, ser, thr;
 (3) acidic: asp, glu;
 (4) basic: asn, gln, his, lys, arg;
 (5) residues that influence chain orientation: gly, pro; and
 40 (6) aromatic: trp, tyr, phe.

Non-conservative substitutions will entail exchanging a member of one of these classes for another class. Such substituted residues also may be introduced into the conservative substitution sites or, more preferably, into the remaining (non-conserved) sites.

45 The variations can be made using methods known in the art such as oligonucleotide-mediated (site-directed) mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed mutagenesis [Carter et al., *Nucl. Acids Res.*, 13:4331 (1986); Zoller et al., *Nucl. Acids Res.*, 10:6487 (1987)], cassette mutagenesis [Wells et al.,

Gene, 34:315 (1985)], restriction selection mutagenesis [Wells et al., Philos. Trans. R. Soc. London Ser. A, 317:415 (1986)] or other known techniques can be performed on the cloned DNA to produce the PRO variant DNA.

Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such 5 amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant [Cunningham and Wells, Science, 244: 1081-1085 (1989)]. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions [Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 10 (1976)]. If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.

C. Modifications of PRO

Covalent modifications of PRO are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a PRO polypeptide with an organic derivatizing 15 agent that is capable of reacting with selected side chains or the N- or C-terminal residues of the PRO. Derivatization with bifunctional agents is useful, for instance, for crosslinking PRO to a water-insoluble support matrix or surface for use in the method for purifying anti-PRO antibodies, and vice-versa. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl 20 esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimidate.

Other modifications include deamidation of glutamyl and asparaginyl residues to the corresponding 25 glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the α -amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the PRO polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence PRO 30 (either by removing the underlying glycosylation site or by deleting the glycosylation by chemical and/or enzymatic means), and/or adding one or more glycosylation sites that are not present in the native sequence PRO. In addition, the phrase includes qualitative changes in the glycosylation of the native proteins, involving a change in the nature and proportions of the various carbohydrate moieties present.

Addition of glycosylation sites to the PRO polypeptide may be accomplished by altering the amino acid 35 sequence. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence PRO (for O-linked glycosylation sites). The PRO amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding

the PRO polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the PRO polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259-306 (1981).

Removal of carbohydrate moieties present on the PRO polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., Arch. Biochem. Biophys., 259:52 (1987) and by Edge et al., Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura et al., Meth. Enzymol., 138:350 (1987).

Another type of covalent modification of PRO comprises linking the PRO polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol (PEG), polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

The PRO of the present invention may also be modified in a way to form a chimeric molecule comprising PRO fused to another, heterologous polypeptide or amino acid sequence.

In one embodiment, such a chimeric molecule comprises a fusion of the PRO with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl- terminus of the PRO. The presence of such epitope-tagged forms of the PRO can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the PRO to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky et al., Protein Engineering, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp et al., BioTechnology, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin et al., Science, 255:192-194 (1992)]; an α -tubulin epitope peptide [Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)].

In an alternative embodiment, the chimeric molecule may comprise a fusion of the PRO with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule (also referred to as an "immunoadhesin"), such a fusion could be to the Fc region of an IgG molecule. The Ig fusions preferably include the substitution of a soluble (transmembrane domain deleted or inactivated) form of a PRO polypeptide in place of at least one variable region within an Ig molecule. In a particularly preferred embodiment, the immunoglobulin fusion includes the hinge, CH2 and CH3, or the hinge, CH1, CH2 and CH3 regions of an IgG1 molecule. For the production of immunoglobulin fusions see also US Patent No. 5,428,130 issued June 27,

1995.

D. Preparation of PRO

The description below relates primarily to production of PRO by culturing cells transformed or transfected with a vector containing PRO nucleic acid. It is, of course, contemplated that alternative methods, 5 which are well known in the art, may be employed to prepare PRO. For instance, the PRO sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [see, e.g., Stewart et al., Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85:2149-2154 (1963)]. *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster 10 City, CA) using manufacturer's instructions. Various portions of the PRO may be chemically synthesized separately and combined using chemical or enzymatic methods to produce the full-length PRO.

1. Isolation of DNA Encoding PRO

DNA encoding PRO may be obtained from a cDNA library prepared from tissue believed to possess the 15 PRO mRNA and to express it at a detectable level. Accordingly, human PRO DNA can be conveniently obtained from a cDNA library prepared from human tissue, such as described in the Examples. The PRO-encoding gene may also be obtained from a genomic library or by known synthetic procedures (e.g., automated nucleic acid synthesis).

Libraries can be screened with probes (such as antibodies to the PRO or oligonucleotides of at least about 20 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding PRO is to use PCR methodology [Sambrook et al., supra; Dieffenbach et al., PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

25 The Examples below describe techniques for screening a cDNA library. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art, and include the use of radiolabels like ³²P-labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high 30 stringency, are provided in Sambrook et al., supra.

Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full-length sequence can be determined using methods known in the art and as described herein.

35 Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures as described in Sambrook et al., supra, to detect precursors and

processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

2. Selection and Transformation of Host Cells

Host cells are transfected or transformed with expression or cloning vectors described herein for PRO production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting 5 transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook et al., supra.

10 Methods of eukaryotic cell transfection and prokaryotic cell transformation are known to the ordinarily skilled artisan, for example, CaCl_2 , CaPO_4 , liposome-mediated and electroporation. Depending on the host cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook et al., supra, or electroporation is generally used for prokaryotes. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as 15 described by Shaw et al., Gene, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transfections have been described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen et al., J. Bact., 130:946 (1977) and Hsiao et al., Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). 20 However, other methods for introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, or polycations, e.g., polybrene, polyornithine, may also be used. For various techniques for transforming mammalian cells, see Keown et al., Methods in Enzymology, 185:527-537 (1990) and Mansour et al., Nature, 336:348-352 (1988).

25 Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited to eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635). Other suitable prokaryotic host cells include Enterobacteriaceae such as *Escherichia*, e.g., *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, 30 e.g., *Salmonella typhimurium*, *Serratia*, e.g., *Serratia marcescens*, and *Shigella*, as well as *Bacilli* such as *B. subtilis* and *B. licheniformis* (e.g., *B. licheniformis* 41P disclosed in DD 266,710 published 12 April 1989), *Pseudomonas* such as *P. aeruginosa*, and *Streptomyces*. These examples are illustrative rather than limiting. Strain W3110 is one particularly preferred host or parent host because it is a common host strain for recombinant DNA product fermentations. Preferably, the host cell secretes minimal amounts of proteolytic enzymes. For 35 example, strain W3110 may be modified to effect a genetic mutation in the genes encoding proteins endogenous to the host, with examples of such hosts including *E. coli* W3110 strain 1A2, which has the complete genotype *tonA*; *E. coli* W3110 strain 9E4, which has the complete genotype *tonA ptr3*; *E. coli* W3110 strain 27C7 (ATCC

55,244), which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT kan'*; *E. coli* W3110 strain 37D6, which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT rbs7 ilvG kan'*; *E. coli* W3110 strain 40B4, which is strain 37D6 with a non-kanamycin resistant *degP* deletion mutation; and an *E. coli* strain having mutant periplasmic protease disclosed in U.S. Patent No. 4,946,783 issued 7 August 1990. Alternatively, *in vitro* methods of cloning, e.g., PCR or other nucleic acid polymerase reactions, are suitable.

5 In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for PRO-encoding vectors. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism. Others include *Schizosaccharomyces pombe* (Beach and Nurse, *Nature*, 290: 140 [1981]; EP 139,383 published 2 May 1985); *Kluyveromyces* hosts (U.S. Patent No. 4,943,529; Fleer et al., *Bio/Technology*, 9:968-975 (1991)) such as, e.g., *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt et al., 10 *J. Bacteriol.*, 154(2):737-742 [1983]), *K. fragilis* (ATCC 12,424), *K. bulgaricus* (ATCC 16,045), *K. wickeramii* (ATCC 24,178), *K. waltii* (ATCC 56,500), *K. drosophilicola* (ATCC 36,906; Van den Berg et al., *Bio/Technology*, 8:135 (1990)), *K. thermotolerans*, and *K. marxianus*; *yarrowiae* (EP 402,226); *Pichia pastoris* (EP 183,070; Sreekrishna et al., *J. Basic Microbiol.*, 28:265-278 [1988]); *Candida*; *Trichoderma reesiae* (EP 244,234); *Neurospora crassa* (Case et al., *Proc. Natl. Acad. Sci. USA*, 76:5259-5263 [1979]); *Schwanniomyces* 15 such as *Schwanniomyces occidentalis* (EP 394,538 published 31 October 1990); and filamentous fungi such as, e.g., *Neurospora*, *Penicillium*, *Tolypocladium* (WO 91/00357 published 10 January 1991), and *Aspergillus* hosts such as *A. nidulans* (Ballance et al., *Biochem. Biophys. Res. Commun.*, 112:284-289 [1983]; Tilburn et al., 20 *Gene*, 26:205-221 [1983]; Yelton et al., *Proc. Natl. Acad. Sci. USA*, 81: 1470-1474 [1984]) and *A. niger* (Kelly and Hynes, *EMBO J.*, 4:475-479 [1985]). Methylotrophic yeasts are suitable herein and include, but are not limited to, yeast capable of growth on methanol selected from the genera consisting of *Hansenula*, *Candida*, *Kloeckera*, *Pichia*, *Saccharomyces*, *Torulopsis*, and *Rhodotorula*. A list of specific species that are exemplary of this class of yeasts may be found in C. Anthony, *The Biochemistry of Methylotrophs*, 269 (1982).

25 Suitable host cells for the expression of glycosylated PRO are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as *Drosophila* S2 and *Spodoptera* Sf9, as well as plant cells. Examples of useful mammalian host cell lines include Chinese hamster ovary (CHO) and COS cells. More specific examples include monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., *J. Gen. Virol.*, 36:59 (1977)); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, *Proc. Natl. Acad. Sci. USA*, 77:4216 (1980)); mouse sertoli cells (TM4, Mather, *Biol. Reprod.*, 23:243-251 (1980)); human lung cells (W138, 30 ATCC CCL 75); human liver cells (Hep G2, HB 8065); and mouse mammary tumor (MMT 060562, ATCC CCL51). The selection of the appropriate host cell is deemed to be within the skill in the art.

3. Selection and Use of a Replicable Vector

35 The nucleic acid (e.g., cDNA or genomic DNA) encoding PRO may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an

appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

5 The PRO may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the PRO-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, 10 penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, e.g., the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* α -factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as 15 signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2μ plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for 20 cloning vectors in mammalian cells.

Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, e.g., ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for *Bacilli*.

25 An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the PRO-encoding nucleic acid, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub et al., *Proc. Natl. Acad. Sci. USA*, 77:4216 (1980). A suitable selection gene for use in yeast is the *trp1* gene present in the yeast plasmid YRp7 [Stinchcomb et al., *Nature*, 282:39 (1979); 30 Kingsman et al., *Gene*, 7:141 (1979); Tschemper et al., *Gene*, 10:157 (1980)]. The *trp1* gene provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1 [Jones, *Genetics*, 85:12 (1977)].

35 Expression and cloning vectors usually contain a promoter operably linked to the PRO-encoding nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the β -lactamase and lactose promoter systems [Chang et al., *Nature*, 275:615 (1978); Goeddel et al., *Nature*, 281:544 (1979)], alkaline phosphatase, a tryptophan (trp) promoter system [Goeddel, *Nucleic Acids Res.*, 8:4057 (1980); EP 36,776], and hybrid

promoters such as the tac promoter [deBoer et al., *Proc. Natl. Acad. Sci. USA*, 80:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding PRO.

5 Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman et al., *J. Biol. Chem.*, 255:2073 (1980)] or other glycolytic enzymes [Hess et al., *J. Adv. Enzyme Reg.*, 7:149 (1968); Holland, *Biochemistry*, 17:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

10 Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytchrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

15 PRO transcription from vectors in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, e.g., the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the host cell systems.

20 Transcription of a DNA encoding the PRO by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α -fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the PRO coding sequence, but is preferably located at a site 5' from the promoter.

30 Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding PRO.

35 Still other methods, vectors, and host cells suitable for adaptation to the synthesis of PRO in recombinant vertebrate cell culture are described in Gething et al., *Nature*, 293:620-625 (1981); Mantei et al., *Nature*, 281:40-46 (1979); EP 117,060; and EP 117,058.

4. Detecting Gene Amplification/Expression

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA [Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)], dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed 5 that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as 10 immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence PRO polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to PRO DNA and encoding a specific 15 antibody epitope.

5. Purification of Polypeptide

Forms of PRO may be recovered from culture medium or from host cell lysates. If membrane-bound, it can be released from the membrane using a suitable detergent solution (e.g. Triton-X 100) or by enzymatic 20 cleavage. Cells employed in expression of PRO can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or cell lysing agents.

It may be desired to purify PRO from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion-exchange column; ethanol 25 precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; protein A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind epitope-tagged forms of the PRO. Various methods of protein purification may be employed and such methods are known in the art and described for example in Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York (1982). The purification step(s) selected will 30 depend, for example, on the nature of the production process used and the particular PRO produced.

E. Uses for PRO

Nucleotide sequences (or their complement) encoding PRO have various applications in the art of molecular biology, including uses as hybridization probes, in chromosome and gene mapping and in the generation 35 of anti-sense RNA and DNA. PRO nucleic acid will also be useful for the preparation of PRO polypeptides by the recombinant techniques described herein.

The full-length native sequence PRO gene, or portions thereof, may be used as hybridization probes for

a cDNA library to isolate the full-length PRO cDNA or to isolate still other cDNAs (for instance, those encoding naturally-occurring variants of PRO or PRO from other species) which have a desired sequence identity to the native PRO sequence disclosed herein. Optionally, the length of the probes will be about 20 to about 50 bases. The hybridization probes may be derived from at least partially novel regions of the full length native nucleotide sequence wherein those regions may be determined without undue experimentation or from genomic sequences including promoters, enhancer elements and introns of native sequence PRO. By way of example, a screening method will comprise isolating the coding region of the PRO gene using the known DNA sequence to synthesize a selected probe of about 40 bases. Hybridization probes may be labeled by a variety of labels, including radionucleotides such as ³²P or ³⁵S, or enzymatic labels such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems. Labeled probes having a sequence complementary to that of the PRO gene of the present invention can be used to screen libraries of human cDNA, genomic DNA or mRNA to determine which members of such libraries the probe hybridizes to. Hybridization techniques are described in further detail in the Examples below.

Any EST sequences disclosed in the present application may similarly be employed as probes, using the methods disclosed herein.

Other useful fragments of the PRO nucleic acids include antisense or sense oligonucleotides comprising a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target PRO mRNA (sense) or PRO DNA (antisense) sequences. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment of the coding region of PRO DNA. Such a fragment generally comprises at least about 14 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, for example, Stein and Cohen (*Cancer Res.* 48:2659, 1988) and van der Krol et al. (*BioTechniques* 6:958, 1988).

Binding of antisense or sense oligonucleotides to target nucleic acid sequences results in the formation of duplexes that block transcription or translation of the target sequence by one of several means, including enhanced degradation of the duplexes, premature termination of transcription or translation, or by other means. The antisense oligonucleotides thus may be used to block expression of PRO proteins. Antisense or sense oligonucleotides further comprise oligonucleotides having modified sugar-phosphodiester backbones (or other sugar linkages, such as those described in WO 91/06629) and wherein such sugar linkages are resistant to endogenous nucleases. Such oligonucleotides with resistant sugar linkages are stable *in vivo* (i.e., capable of resisting enzymatic degradation) but retain sequence specificity to be able to bind to target nucleotide sequences.

Other examples of sense or antisense oligonucleotides include those oligonucleotides which are covalently linked to organic moieties, such as those described in WO 90/10048, and other moieties that increase affinity of the oligonucleotide for a target nucleic acid sequence, such as poly-(L-lysine). Further still, intercalating agents, such as ellipticine, and alkylating agents or metal complexes may be attached to sense or antisense oligonucleotides to modify binding specificities of the antisense or sense oligonucleotide for the target nucleotide sequence.

Antisense or sense oligonucleotides may be introduced into a cell containing the target nucleic acid sequence by any gene transfer method, including, for example, CaPO₄-mediated DNA transfection,

electroporation, or by using gene transfer vectors such as Epstein-Barr virus. In a preferred procedure, an antisense or sense oligonucleotide is inserted into a suitable retroviral vector. A cell containing the target nucleic acid sequence is contacted with the recombinant retroviral vector, either *in vivo* or *ex vivo*. Suitable retroviral vectors include, but are not limited to, those derived from the murine retrovirus M-MuLV, N2 (a retrovirus derived from M-MuLV), or the double copy vectors designated DCT5A, DCT5B and DCT5C (see WO 5 90/13641).

Sense or antisense oligonucleotides also may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does 10 not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell.

Alternatively, a sense or an antisense oligonucleotide may be introduced into a cell containing the target nucleic acid sequence by formation of an oligonucleotide-lipid complex, as described in WO 90/10448. The sense or antisense oligonucleotide-lipid complex is preferably dissociated within the cell by an endogenous lipase.

15 Antisense or sense RNA or DNA molecules are generally at least about 5 bases in length, about 10 bases in length, about 15 bases in length, about 20 bases in length, about 25 bases in length, about 30 bases in length, about 35 bases in length, about 40 bases in length, about 45 bases in length, about 50 bases in length, about 55 bases in length, about 60 bases in length, about 65 bases in length, about 70 bases in length, about 75 bases in length, about 80 bases in length, about 85 bases in length, about 90 bases in length, about 95 bases in length, 20 about 100 bases in length, or more.

The probes may also be employed in PCR techniques to generate a pool of sequences for identification of closely related PRO coding sequences.

25 Nucleotide sequences encoding a PRO can also be used to construct hybridization probes for mapping the gene which encodes that PRO and for the genetic analysis of individuals with genetic disorders. The nucleotide sequences provided herein may be mapped to a chromosome and specific regions of a chromosome using known techniques, such as *in situ* hybridization, linkage analysis against known chromosomal markers, and hybridization screening with libraries.

30 When the coding sequences for PRO encode a protein which binds to another protein (example, where the PRO is a receptor), the PRO can be used in assays to identify the other proteins or molecules involved in the binding interaction. By such methods, inhibitors of the receptor/ligand binding interaction can be identified. Proteins involved in such binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction. Also, the receptor PRO can be used to isolate correlative ligand(s). Screening assays can be designed to find lead compounds that mimic the biological activity of a native PRO or a receptor for PRO. Such screening assays will include assays amenable to high-throughput screening of chemical 35 libraries, making them particularly suitable for identifying small molecule drug candidates. Small molecules contemplated include synthetic organic or inorganic compounds. The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays and cell based

assays, which are well characterized in the art.

Nucleic acids which encode PRO or its modified forms can also be used to generate either transgenic animals or "knock out" animals which, in turn, are useful in the development and screening of therapeutically useful reagents. A transgenic animal (e.g., a mouse or rat) is an animal having cells that contain a transgene, which transgene was introduced into the animal or an ancestor of the animal at a prenatal, e.g., an embryonic stage. A transgene is a DNA which is integrated into the genome of a cell from which a transgenic animal develops. In one embodiment, cDNA encoding PRO can be used to clone genomic DNA encoding PRO in accordance with established techniques and the genomic sequences used to generate transgenic animals that contain cells which express DNA encoding PRO. Methods for generating transgenic animals, particularly animals such as mice or rats, have become conventional in the art and are described, for example, in U.S. Patent Nos: 4,736,866 and 4,870,009. Typically, particular cells would be targeted for PRO transgene incorporation with tissue-specific enhancers. Transgenic animals that include a copy of a transgene encoding PRO introduced into the germ line of the animal at an embryonic stage can be used to examine the effect of increased expression of DNA encoding PRO. Such animals can be used as tester animals for reagents thought to confer protection from, for example, pathological conditions associated with its overexpression. In accordance with this facet of the invention, an animal is treated with the reagent and a reduced incidence of the pathological condition, compared to untreated animals bearing the transgene, would indicate a potential therapeutic intervention for the pathological condition.

Alternatively, non-human homologues of PRO can be used to construct a PRO "knock out" animal which has a defective or altered gene encoding PRO as a result of homologous recombination between the endogenous gene encoding PRO and altered genomic DNA encoding PRO introduced into an embryonic stem cell of the animal. For example, cDNA encoding PRO can be used to clone genomic DNA encoding PRO in accordance with established techniques. A portion of the genomic DNA encoding PRO can be deleted or replaced with another gene, such as a gene encoding a selectable marker which can be used to monitor integration. Typically, several kilobases of unaltered flanking DNA (both at the 5' and 3' ends) are included in the vector [see e.g., Thomas and Capecchi, *Cell*, 51:503 (1987) for a description of homologous recombination vectors]. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced DNA has homologously recombined with the endogenous DNA are selected [see e.g., Li et al., *Cell*, 69:915 (1992)]. The selected cells are then injected into a blastocyst of an animal (e.g., a mouse or rat) to form aggregation chimeras [see e.g., Bradley, in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson, ed. (IRL, Oxford, 1987), pp. 113-152]. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term to create a "knock out" animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knockout animals can be characterized for instance, for their ability to defend against certain pathological conditions and for their development of pathological conditions due to absence of the PRO polypeptide.

Nucleic acid encoding the PRO polypeptides may also be used in gene therapy. In gene therapy applications, genes are introduced into cells in order to achieve *in vivo* synthesis of a therapeutically effective

genetic product, for example for replacement of a defective gene. "Gene therapy" includes both conventional gene therapy where a lasting effect is achieved by a single treatment, and the administration of gene therapeutic agents, which involves the one time or repeated administration of a therapeutically effective DNA or mRNA. Antisense RNAs and DNAs can be used as therapeutic agents for blocking the expression of certain genes *in vivo*. It has already been shown that short antisense oligonucleotides can be imported into cells where they act as 5 inhibitors, despite their low intracellular concentrations caused by their restricted uptake by the cell membrane. (Zamecnik *et al.*, Proc. Natl. Acad. Sci. USA 83:4143-4146 [1986]). The oligonucleotides can be modified to enhance their uptake, e.g. by substituting their negatively charged phosphodiester groups by uncharged groups.

There are a variety of techniques available for introducing nucleic acids into viable cells. The techniques vary depending upon whether the nucleic acid is transferred into cultured cells *in vitro*, or *in vivo* in the cells of 10 the intended host. Techniques suitable for the transfer of nucleic acid into mammalian cells *in vitro* include the use of liposomes, electroporation, microinjection, cell fusion, DEAE-dextran, the calcium phosphate precipitation method, etc. The currently preferred *in vivo* gene transfer techniques include transfection with viral (typically retroviral) vectors and viral coat protein-liposome mediated transfection (Dzau *et al.*, Trends in Biotechnology 11, 205-210 [1993]). In some situations it is desirable to provide the nucleic acid source with an agent that targets 15 the target cells, such as an antibody specific for a cell surface membrane protein or the target cell, a ligand for a receptor on the target cell, etc. Where liposomes are employed, proteins which bind to a cell surface membrane protein associated with endocytosis may be used for targeting and/or to facilitate uptake, e.g. capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellular localization and enhance intracellular half-life. The technique of receptor-mediated endocytosis is described, for example, by Wu *et al.*, J. Biol. Chem. 262, 4429-4432 (1987); and Wagner 20 *et al.*, Proc. Natl. Acad. Sci. USA 87, 3410-3414 (1990). For review of gene marking and gene therapy protocols see Anderson *et al.*, Science 256, 808-813 (1992).

The PRO polypeptides described herein may also be employed as molecular weight markers for protein 25 electrophoresis purposes and the isolated nucleic acid sequences may be used for recombinantly expressing those markers.

The nucleic acid molecules encoding the PRO polypeptides or fragments thereof described herein are useful for chromosome identification. In this regard, there exists an ongoing need to identify new chromosome markers, since relatively few chromosome marking reagents, based upon actual sequence data are presently available. Each PRO nucleic acid molecule of the present invention can be used as a chromosome marker.

30 The PRO polypeptides and nucleic acid molecules of the present invention may also be used diagnostically for tissue typing, wherein the PRO polypeptides of the present invention may be differentially expressed in one tissue as compared to another, preferably in a diseased tissue as compared to a normal tissue of the same tissue type. PRO nucleic acid molecules will find use for generating probes for PCR, Northern analysis, Southern analysis and Western analysis.

35 The PRO polypeptides described herein may also be employed as therapeutic agents. The PRO polypeptides of the present invention can be formulated according to known methods to prepare pharmaceutically useful compositions, whereby the PRO product hereof is combined in admixture with a pharmaceutically

acceptable carrier vehicle. Therapeutic formulations are prepared for storage by mixing the active ingredient having the desired degree of purity with optional physiologically acceptable carriers, excipients or stabilizers (Remington's Pharmaceutical Sciences 16th edition, Osol, A. Ed. (1980)), in the form of lyophilized formulations or aqueous solutions. Acceptable carriers, excipients or stabilizers are nontoxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptides; proteins, such as serum albumin, gelatin or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone, amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides and other carbohydrates including glucose, mannose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN™, PLURONICS™ or PEG.

5 The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes, prior to or following lyophilization and reconstitution.

Therapeutic compositions herein generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

10 The route of administration is in accord with known methods, e.g. injection or infusion by intravenous, intraperitoneal, intracerebral, intramuscular, intraocular, intraarterial or intralesional routes, topical administration, or by sustained release systems.

15 Dosages and desired drug concentrations of pharmaceutical compositions of the present invention may vary depending on the particular use envisioned. The determination of the appropriate dosage or route of administration is well within the skill of an ordinary physician. Animal experiments provide reliable guidance for the determination of effective doses for human therapy. Interspecies scaling of effective doses can be performed following the principles laid down by Mordini, J. and Chappell, W. "The use of interspecies scaling in toxicokinetics" In Toxicokinetics and New Drug Development, Yacobi et al., Eds., Pergamon Press, New York 1989, pp. 42-96.

20 When *in vivo* administration of a PRO polypeptide or agonist or antagonist thereof is employed, normal dosage amounts may vary from about 10 ng/kg to up to 100 mg/kg of mammal body weight or more per day, preferably about 1 μ g/kg/day to 10 mg/kg/day, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature; see, for example, U.S. Pat. Nos. 4,657,760; 5,206,344; or 5,225,212. It is anticipated that different formulations will be effective for different 25 treatment compounds and different disorders, that administration targeting one organ or tissue, for example, may necessitate delivery in a manner different from that to another organ or tissue.

30 Where sustained-release administration of a PRO polypeptide is desired in a formulation with release characteristics suitable for the treatment of any disease or disorder requiring administration of the PRO polypeptide, microencapsulation of the PRO polypeptide is contemplated. Microencapsulation of recombinant proteins for sustained release has been successfully performed with human growth hormone (rhGH), interferon- (rhIFN-), interleukin-2, and MN rgp120. Johnson et al., Nat. Med., 2:795-799 (1996); Yasuda, Biomed. Ther., 27:1221-1223 (1993); Hora et al., Bio/Technology, 8:755-758 (1990); Cleland, "Design and Production of Single

Immunization Vaccines Using Polylactide Polyglycolide Microsphere Systems," in Vaccine Design: The Subunit and Adjuvant Approach, Powell and Newman, eds, (Plenum Press: New York, 1995), pp. 439-462; WO 97/03692, WO 96/40072, WO 96/07399; and U.S. Pat. No. 5,654,010.

15 The sustained-release formulations of these proteins were developed using poly-lactic-glycolic acid (PLGA) polymer due to its biocompatibility and wide range of biodegradable properties. The degradation products of PLGA, lactic and glycolic acids, can be cleared quickly within the human body. Moreover, the degradability of this polymer can be adjusted from months to years depending on its molecular weight and composition. Lewis, "Controlled release of bioactive agents from lactide/glycolide polymer," in: M. Chasin and R. Langer (Eds.), Biodegradable Polymers as Drug Delivery Systems (Marcel Dekker: New York, 1990), pp. 1-41.

10 This invention encompasses methods of screening compounds to identify those that mimic the PRO polypeptide (agonists) or prevent the effect of the PRO polypeptide (antagonists). Screening assays for antagonist drug candidates are designed to identify compounds that bind or complex with the PRO polypeptides encoded by the genes identified herein, or otherwise interfere with the interaction of the encoded polypeptides with other cellular proteins. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates.

15 The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays, and cell-based assays, which are well characterized in the art.

20 All assays for antagonists are common in that they call for contacting the drug candidate with a PRO polypeptide encoded by a nucleic acid identified herein under conditions and for a time sufficient to allow these two components to interact.

25 In binding assays, the interaction is binding and the complex formed can be isolated or detected in the reaction mixture. In a particular embodiment, the PRO polypeptide encoded by the gene identified herein or the drug candidate is immobilized on a solid phase, e.g., on a microtiter plate, by covalent or non-covalent attachments. Non-covalent attachment generally is accomplished by coating the solid surface with a solution of the PRO polypeptide and drying. Alternatively, an immobilized antibody, e.g., a monoclonal antibody, specific for the PRO polypeptide to be immobilized can be used to anchor it to a solid surface. The assay is performed by adding the non-immobilized component, which may be labeled by a detectable label, to the immobilized component, e.g., the coated surface containing the anchored component. When the reaction is complete, the non-reacted components are removed, e.g., by washing, and complexes anchored on the solid surface are detected. When the originally non-immobilized component carries a detectable label, the detection of label immobilized on 30 the surface indicates that complexing occurred. Where the originally non-immobilized component does not carry a label, complexing can be detected, for example, by using a labeled antibody specifically binding the immobilized complex.

35 If the candidate compound interacts with but does not bind to a particular PRO polypeptide encoded by a gene identified herein, its interaction with that polypeptide can be assayed by methods well known for detecting protein-protein interactions. Such assays include traditional approaches, such as, e.g., cross-linking, co-immunoprecipitation, and co-purification through gradients or chromatographic columns. In addition, protein-protein interactions can be monitored by using a yeast-based genetic system described by Fields and co-workers

(Fields and Song, Nature (London), 340:245-246 (1989); Chien et al., Proc. Natl. Acad. Sci. USA, 88:9578-9582 (1991)) as disclosed by Chevray and Nathans, Proc. Natl. Acad. Sci. USA, 89: 5789-5793 (1991). Many transcriptional activators, such as yeast GAL4, consist of two physically discrete modular domains, one acting as the DNA-binding domain, the other one functioning as the transcription-activation domain. The yeast expression system described in the foregoing publications (generally referred to as the "two-hybrid system") takes 5 advantage of this property, and employs two hybrid proteins, one in which the target protein is fused to the DNA-binding domain of GAL4, and another, in which candidate activating proteins are fused to the activation domain. The expression of a GAL1-lacZ reporter gene under control of a GAL4-activated promoter depends on reconstitution of GAL4 activity via protein-protein interaction. Colonies containing interacting polypeptides are detected with a chromogenic substrate for β -galactosidase. A complete kit (MATCHMAKERTM) for identifying 10 protein-protein interactions between two specific proteins using the two-hybrid technique is commercially available from Clontech. This system can also be extended to map protein domains involved in specific protein interactions as well as to pinpoint amino acid residues that are crucial for these interactions.

Compounds that interfere with the interaction of a gene encoding a PRO polypeptide identified herein and other intra- or extracellular components can be tested as follows: usually a reaction mixture is prepared 15 containing the product of the gene and the intra- or extracellular component under conditions and for a time allowing for the interaction and binding of the two products. To test the ability of a candidate compound to inhibit binding, the reaction is run in the absence and in the presence of the test compound. In addition, a placebo may be added to a third reaction mixture, to serve as positive control. The binding (complex formation) between the test compound and the intra- or extracellular component present in the mixture is monitored as described 20 hereinabove. The formation of a complex in the control reaction(s) but not in the reaction mixture containing the test compound indicates that the test compound interferes with the interaction of the test compound and its reaction partner.

To assay for antagonists, the PRO polypeptide may be added to a cell along with the compound to be screened for a particular activity and the ability of the compound to inhibit the activity of interest in the presence 25 of the PRO polypeptide indicates that the compound is an antagonist to the PRO polypeptide. Alternatively, antagonists may be detected by combining the PRO polypeptide and a potential antagonist with membrane-bound PRO polypeptide receptors or recombinant receptors under appropriate conditions for a competitive inhibition assay. The PRO polypeptide can be labeled, such as by radioactivity, such that the number of PRO polypeptide molecules bound to the receptor can be used to determine the effectiveness of the potential antagonist. The gene 30 encoding the receptor can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting. Coligan et al., Current Protocols in Immun., 1(2): Chapter 5 (1991). Preferably, expression cloning is employed wherein polyadenylated RNA is prepared from a cell responsive to the PRO polypeptide and a cDNA library created from this RNA is divided into pools and used to transfect COS cells or other cells that are not responsive to the PRO polypeptide. Transfected cells that are grown on glass 35 slides are exposed to labeled PRO polypeptide. The PRO polypeptide can be labeled by a variety of means including iodination or inclusion of a recognition site for a site-specific protein kinase. Following fixation and incubation, the slides are subjected to autoradiographic analysis. Positive pools are identified and sub-pools are

prepared and re-transfected using an interactive sub-pooling and re-screening process, eventually yielding a single clone that encodes the putative receptor.

As an alternative approach for receptor identification, labeled PRO polypeptide can be photoaffinity-linked with cell membrane or extract preparations that express the receptor molecule. Cross-linked material is resolved by PAGE and exposed to X-ray film. The labeled complex containing the receptor can be excised, 5 resolved into peptide fragments, and subjected to protein micro-sequencing. The amino acid sequence obtained from micro- sequencing would be used to design a set of degenerate oligonucleotide probes to screen a cDNA library to identify the gene encoding the putative receptor.

In another assay for antagonists, mammalian cells or a membrane preparation expressing the receptor would be incubated with labeled PRO polypeptide in the presence of the candidate compound. The ability of the 10 compound to enhance or block this interaction could then be measured.

More specific examples of potential antagonists include an oligonucleotide that binds to the fusions of immunoglobulin with PRO polypeptide, and, in particular, antibodies including, without limitation, poly- and monoclonal antibodies and antibody fragments, single-chain antibodies, anti-idiotypic antibodies, and chimeric or humanized versions of such antibodies or fragments, as well as human antibodies and antibody fragments. 15 Alternatively, a potential antagonist may be a closely related protein, for example, a mutated form of the PRO polypeptide that recognizes the receptor but imparts no effect, thereby competitively inhibiting the action of the PRO polypeptide.

Another potential PRO polypeptide antagonist is an antisense RNA or DNA construct prepared using antisense technology, where, e.g., an antisense RNA or DNA molecule acts to block directly the translation of 20 mRNA by hybridizing to targeted mRNA and preventing protein translation. Antisense technology can be used to control gene expression through triple-helix formation or antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion of the polynucleotide sequence, which encodes the mature PRO polypeptides herein, is used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region 25 of the gene involved in transcription (triple helix - see Lee et al., *Nucl. Acids Res.*, 6:3073 (1979); Cooney et al., *Science*, 241: 456 (1988); Dervan et al., *Science*, 251:1360 (1991)), thereby preventing transcription and the production of the PRO polypeptide. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into the PRO polypeptide (antisense - Okano, *Neurochem.*, 56:560 (1991); *Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression* (CRC Press: Boca Raton, FL, 1988). 30 The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of the PRO polypeptide. When antisense DNA is used, oligodeoxyribonucleotides derived from the translation-initiation site, e.g., between about -10 and +10 positions of the target gene nucleotide sequence, are preferred.

Potential antagonists include small molecules that bind to the active site, the receptor binding site, or 35 growth factor or other relevant binding site of the PRO polypeptide, thereby blocking the normal biological activity of the PRO polypeptide. Examples of small molecules include, but are not limited to, small peptides or peptide-like molecules, preferably soluble peptides, and synthetic non-peptidyl organic or inorganic compounds.

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. Ribozymes act by sequence-specific hybridization to the complementary target RNA, followed by endonucleolytic cleavage. Specific ribozyme cleavage sites within a potential RNA target can be identified by known techniques. For further details see, e.g., Rossi, *Current Biology*, 4:469-471 (1994), and PCT publication No. WO 97/33551 (published September 18, 1997).

5 Nucleic acid molecules in triple-helix formation used to inhibit transcription should be single-stranded and composed of deoxynucleotides. The base composition of these oligonucleotides is designed such that it promotes triple-helix formation via Hoogsteen base-pairing rules, which generally require sizeable stretches of purines or pyrimidines on one strand of a duplex. For further details see, e.g., PCT publication No. WO 97/33551, *supra*.

10 These small molecules can be identified by any one or more of the screening assays discussed hereinabove and/or by any other screening techniques well known for those skilled in the art.

Diagnostic and therapeutic uses of the herein disclosed molecules may also be based upon the positive functional assay hits disclosed and described below.

15 F. Anti-PRO Antibodies

The present invention further provides anti-PRO antibodies. Exemplary antibodies include polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies.

1. Polyclonal Antibodies

20 The anti-PRO antibodies may comprise polyclonal antibodies. Methods of preparing polyclonal antibodies are known to the skilled artisan. Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include the PRO polypeptide or a fusion protein thereof. It may be useful to conjugate 25 the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

30

2. Monoclonal Antibodies

The anti-PRO antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein, *Nature*, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an 35 immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized *in vitro*.

The immunizing agent will typically include the PRO polypeptide or a fusion protein thereof. Generally,

either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell [Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103]. Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies [Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63].

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against PRO. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980).

After the desired hybridoma cells are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods [Goding, supra]. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells may be grown *in vivo* as ascites in a mammal.

The monoclonal antibodies secreted by the subclones may be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal

antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences [U.S. Patent No. 4,816,567; Morrison et al., *supra*] or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

5 The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is truncated generally at any point in the Fc region so as to prevent heavy chain crosslinking. Alternatively, the relevant cysteine residues are substituted with another amino acid residue 10 or are deleted so as to prevent crosslinking.

10 *In vitro* methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art.

15 3. Human and Humanized Antibodies

The anti-PRO antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin.

20 Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor

25 in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones et al., 30 *Nature*, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-329 (1988); and Presta, *Curr. Op. Struct. Biol.*, 2:593-596 (1992)].

35 Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and co-workers [Jones et al., *Nature*, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-327 (1988); Verhoeyen et al., *Science*, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR sequences for the corresponding sequences

of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

5 Human antibodies can also be produced using various techniques known in the art, including phage display libraries [Hoogenboom and Winter, *J. Mol. Biol.*, 227:381 (1991); Marks et al., *J. Mol. Biol.*, 222:581 (1991)]. The techniques of Cole et al. and Boerner et al. are also available for the preparation of human monoclonal antibodies (Cole et al., *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, p. 77 (1985) and Boerner et al., *J. Immunol.*, 147(1):86-95 (1991)]. Similarly, human antibodies can be made by introducing of 10 human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in the following scientific publications: Marks et al., *Bio/Technology* 10, 15 779-783 (1992); Lonberg et al., *Nature* 368 856-859 (1994); Morrison, *Nature* 368, 812-13 (1994); Fishwild et al., *Nature Biotechnology* 14, 845-51 (1996); Neuberger, *Nature Biotechnology* 14, 826 (1996); Lonberg and Huszar, *Intern. Rev. Immunol.* 13 65-93 (1995).

20 The antibodies may also be affinity matured using known selection and/or mutagenesis methods as described above. Preferred affinity matured antibodies have an affinity which is five times, more preferably 10 times, even more preferably 20 or 30 times greater than the starting antibody (generally murine, humanized or human) from which the matured antibody is prepared.

4. Bispecific Antibodies

25 Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for the PRO, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit.

30 Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities [Milstein and Cuello, *Nature*, 305:537-539 (1983)]. Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker et al., *EMBO J.*, 10:3655-3659 (1991).

35 Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the

first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

5 According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar 10 size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

15 Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(ab')₂, bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate F(ab')₂ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives 20 is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

25 Fab' fragments may be directly recovered from *E. coli* and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody F(ab')₂ molecule. Each Fab' fragment was separately secreted from *E. coli* and subjected to directed chemical coupling *in vitro* to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

30 Various technique for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by 35 Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on

the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber *et al.*, *J. Immunol.* 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared.

5 Tutt *et al.*, *J. Immunol.* 147:60 (1991).

Exemplary bispecific antibodies may bind to two different epitopes on a given PRO polypeptide herein. Alternatively, an anti-PRO polypeptide arm may be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (Fc γ R), such as Fc γ RI (CD64), Fc γ RII (CD32) and Fc γ RIII (CD16) so as to focus cellular defense mechanisms 10 to the cell expressing the particular PRO polypeptide. Bispecific antibodies may also be used to localize cytotoxic agents to cells which express a particular PRO polypeptide. These antibodies possess a PRO-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the PRO polypeptide and further binds tissue factor (TF).

15 5. Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells [U.S. Patent No. 4,676,980], and for treatment of HIV infection [WO 91/00360; WO 92/200373; EP 03089]. It is contemplated that the antibodies may be prepared *in vitro* using 20 known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptoputyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

25 6. Effector Function Engineering

It may be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, *e.g.*, the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) may be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated may have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron *et al.*, *J. Exp Med.*, 176: 1191-1195 (1992) and Shope, *J. Immunol.*, 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity may also be prepared using heterobifunctional cross-linkers as described in Wolff *et al.* *Cancer Research*, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and may thereby have enhanced complement lysis and ADCC capabilities. See Stevenson *et al.*, *Anti-Cancer Drug Design*, 35 3: 219-230 (1989).

7. Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active 5 fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolaca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the trichothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ^{212}Bi , ^{131}I , ^{131}In , ^{90}Y , and ^{186}Re .

10 Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridylthiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-15 active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta *et al.*, *Science*, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody may be conjugated to a "receptor" (such streptavidin) for utilization 20 in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is conjugated to a cytotoxic agent (e.g., a radionucleotide).

8. Immunoliposomes

25 The antibodies disclosed herein may also be formulated as immunoliposomes. Liposomes containing the antibody are prepared by methods known in the art, such as described in Epstein *et al.*, *Proc. Natl. Acad. Sci. USA*, 82: 3688 (1985); Hwang *et al.*, *Proc. Natl. Acad. Sci. USA*, 77: 4030 (1980); and U.S. Pat. Nos. 4,485,045 and 4,544,545. Liposomes with enhanced circulation time are disclosed in U.S. Patent No. 5,013,556.

30 Particularly useful liposomes can be generated by the reverse-phase evaporation method with a lipid composition comprising phosphatidylcholine, cholesterol, and PEG-derivatized phosphatidylethanolamine (PEG-PE). Liposomes are extruded through filters of defined pore size to yield liposomes with the desired diameter. Fab' fragments of the antibody of the present invention can be conjugated to the liposomes as described in Martin *et al.*, *J. Biol. Chem.*, 257: 286-288 (1982) via a disulfide-interchange reaction. A chemotherapeutic agent (such as Doxorubicin) is optionally contained within the liposome. See Gabizon *et al.*, *J. National Cancer Inst.*, 81(19): 35 1484 (1989).

9. Pharmaceutical Compositions of Antibodies

Antibodies specifically binding a PRO polypeptide identified herein, as well as other molecules identified by the screening assays disclosed hereinbefore, can be administered for the treatment of various disorders in the form of pharmaceutical compositions.

If the PRO polypeptide is intracellular and whole antibodies are used as inhibitors, internalizing antibodies are preferred. However, lipofections or liposomes can also be used to deliver the antibody, or an antibody fragment, into cells. Where antibody fragments are used, the smallest inhibitory fragment that specifically binds to the binding domain of the target protein is preferred. For example, based upon the variable-region sequences of an antibody, peptide molecules can be designed that retain the ability to bind the target protein sequence. Such peptides can be synthesized chemically and/or produced by recombinant DNA technology. See, 10 *e.g.*, Marasco *et al.*, Proc. Natl. Acad. Sci. USA, 90: 7889-7893 (1993). The formulation herein may also contain more than one active compound as necessary for the particular indication being treated, preferably those with complementary activities that do not adversely affect each other. Alternatively, or in addition, the composition may comprise an agent that enhances its function, such as, for example, a cytotoxic agent, cytokine, 15 chemotherapeutic agent, or growth-inhibitory agent. Such molecules are suitably present in combination in amounts that are effective for the purpose intended.

The active ingredients may also be entrapped in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization, for example, hydroxymethylcellulose or gelatin-microcapsules and poly-(methylmethacrylate) microcapsules, respectively, in colloidal drug delivery systems (for example, liposomes, 20 albumin microspheres, microemulsions, nano-particles, and nanocapsules) or in macroemulsions. Such techniques are disclosed in Remington's Pharmaceutical Sciences, supra.

The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes.

Sustained-release preparations may be prepared. Suitable examples of sustained-release preparations include semipermeable matrices of solid hydrophobic polymers containing the antibody, which matrices are in 25 the form of shaped articles, *e.g.*, films, or microcapsules. Examples of sustained-release matrices include polyesters, hydrogels (for example, poly(2-hydroxyethyl-methacrylate), or poly(vinylalcohol)), polylactides (U.S. Pat. No. 3,773,919), copolymers of L-glutamic acid and γ ethyl-L-glutamate, non-degradable ethylene-vinyl

acetate, degradable lactic acid-glycolic acid copolymers such as the LUPRON DEPOTTM (injectable microspheres composed of lactic acid-glycolic acid copolymer and leuprolide acetate), and poly-D-(-)-3-hydroxybutyric acid.

30 While polymers such as ethylene-vinyl acetate and lactic acid-glycolic acid enable release of molecules for over 100 days, certain hydrogels release proteins for shorter time periods. When encapsulated antibodies remain in the body for a long time, they may denature or aggregate as a result of exposure to moisture at 37°C, resulting in a loss of biological activity and possible changes in immunogenicity. Rational strategies can be devised for stabilization depending on the mechanism involved. For example, if the aggregation mechanism is discovered

35 to be intermolecular S-S bond formation through thio-disulfide interchange, stabilization may be achieved by modifying sulphydryl residues, lyophilizing from acidic solutions, controlling moisture content, using appropriate additives, and developing specific polymer matrix compositions.

G. Uses for anti-PRO Antibodies

The anti-PRO antibodies of the invention have various utilities. For example, anti-PRO antibodies may be used in diagnostic assays for PRO, *e.g.*, detecting its expression (and in some cases, differential expression) in specific cells, tissues, or serum. Various diagnostic assay techniques known in the art may be used, such as competitive binding assays, direct or indirect sandwich assays and immunoprecipitation assays conducted in either 5 heterogeneous or homogeneous phases [Zola, Monoclonal Antibodies: A Manual of Techniques, CRC Press, Inc. (1987) pp. 147-158]. The antibodies used in the diagnostic assays can be labeled with a detectable moiety. The detectable moiety should be capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as ^3H , ^{14}C , ^{32}P , ^{35}S , or ^{125}I , a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline 10 phosphatase, beta-galactosidase or horseradish peroxidase. Any method known in the art for conjugating the antibody to the detectable moiety may be employed, including those methods described by Hunter et al., Nature, 144:945 (1962); David et al., Biochemistry, 13:1014 (1974); Pain et al., J. Immunol. Meth., 40:219 (1981); and Nygren, J. Histochem. and Cytochem., 30:407 (1982).

Anti-PRO antibodies also are useful for the affinity purification of PRO from recombinant cell culture 15 or natural sources. In this process, the antibodies against PRO are immobilized on a suitable support, such a Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody then is contacted with a sample containing the PRO to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except the PRO, which is bound to the immobilized antibody. Finally, the support is washed with another suitable solvent that will release the PRO from the 20 antibody.

The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

25

EXAMPLES

Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cells identified in the following examples, and throughout the specification, by ATCC accession numbers is the American Type Culture Collection, Manassas, 30 VA.

EXAMPLE 1: Extracellular Domain Homology Screening to Identify Novel Polypeptides and cDNA Encoding Therefor

The extracellular domain (ECD) sequences (including the secretion signal sequence, if any) from about 35 950 known secreted proteins from the Swiss-Prot public database were used to search EST databases. The EST databases included public databases (*e.g.*, Dayhoff, GenBank), and proprietary databases (*e.g.* LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or

BLAST-2 (Altschul *et al.*, Methods in Enzymology, 266:460-480 (1996)) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequences. Those comparisons with a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, WA).

Using this extracellular domain homology screen, consensus DNA sequences were assembled relative 5 to the other identified EST sequences using phrap. In addition, the consensus DNA sequences obtained were often (but not always) extended using repeated cycles of BLAST or BLAST-2 and phrap to extend the consensus sequence as far as possible using the sources of EST sequences discussed above.

Based upon the consensus sequences obtained as described above, oligonucleotides were then synthesized and used to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to 10 isolate a clone of the full-length coding sequence for a PRO polypeptide. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5kbp. In order to screen several libraries for 15 a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel *et al.*, Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with 20 oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRK5B or pRK5D; pRK5B is a precursor of pRK5D that does not contain the SfiI site; *see*, Holmes *et al.*, Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

EXAMPLE 2: Isolation of cDNA clones by Amylase Screening

25 1. Preparation of oligo dT primed cDNA library

mRNA was isolated from a human tissue of interest using reagents and protocols from Invitrogen, San 30 Diego, CA (Fast Track 2). This RNA was used to generate an oligo dT primed cDNA library in the vector pRK5D using reagents and protocols from Life Technologies, Gaithersburg, MD (Super Script Plasmid System). In this procedure, the double stranded cDNA was sized to greater than 1000 bp and the SalI/NotI linker cDNA was cloned into XhoI/NotI cleaved vector. pRK5D is a cloning vector that has an sp6 transcription initiation site followed by an SfiI restriction enzyme site preceding the XhoI/NotI cDNA cloning sites.

2. Preparation of random primed cDNA library

A secondary cDNA library was generated in order to preferentially represent the 5' ends of the primary 35 cDNA clones. Sp6 RNA was generated from the primary library (described above), and this RNA was used to generate a random primed cDNA library in the vector pSST-AMY.0 using reagents and protocols from Life Technologies (Super Script Plasmid System, referenced above). In this procedure the double stranded cDNA was

sized to 500-1000 bp, linkerered with blunt to *NotI* adaptors, cleaved with *SfiI*, and cloned into *SfiI/NotI* cleaved vector. pSST-AMY.0 is a cloning vector that has a yeast alcohol dehydrogenase promoter preceding the cDNA cloning sites and the mouse amylase sequence (the mature sequence without the secretion signal) followed by the yeast alcohol dehydrogenase terminator, after the cloning sites. Thus, cDNAs cloned into this vector that are fused in frame with amylase sequence will lead to the secretion of amylase from appropriately transfected yeast colonies.

5

3. Transformation and Detection

DNA from the library described in paragraph 2 above was chilled on ice to which was added electrocompetent DH10B bacteria (Life Technologies, 20 ml). The bacteria and vector mixture was then 10 electroporated as recommended by the manufacturer. Subsequently, SOC media (Life Technologies, 1 ml) was added and the mixture was incubated at 37°C for 30 minutes. The transformants were then plated onto 20 standard 150 mm LB plates containing ampicillin and incubated for 16 hours (37°C). Positive colonies were scraped off the plates and the DNA was isolated from the bacterial pellet using standard protocols, *e.g.* CsCl-gradient. The purified DNA was then carried on to the yeast protocols below.

15

The yeast methods were divided into three categories: (1) Transformation of yeast with the plasmid/cDNA combined vector; (2) Detection and isolation of yeast clones secreting amylase; and (3) PCR amplification of the insert directly from the yeast colony and purification of the DNA for sequencing and further analysis.

20

The yeast strain used was HD56-5A (ATCC-90785). This strain has the following genotype: MAT alpha, *ura3-52*, *leu2-3*, *leu2-112*, *his3-11*, *his3-15*, *MAL*⁺, *SUC*⁺, *GAL*⁺. Preferably, yeast mutants can be employed that have deficient post-translational pathways. Such mutants may have translocation deficient alleles in *sec71*, *sec72*, *sec62*, with truncated *sec71* being most preferred. Alternatively, antagonists (including antisense nucleotides and/or ligands) which interfere with the normal operation of these genes, other proteins implicated in this post translation pathway (*e.g.*, *SEC61p*, *SEC72p*, *SEC62p*, *SEC63p*, *TDJ1p* or *SSA1p-4p*) or the complex formation of these proteins may also be preferably employed in combination with the amylase-expressing yeast.

25

Transformation was performed based on the protocol outlined by Gietz *et al.*, *Nucl. Acid. Res.*, 20:1425 (1992). Transformed cells were then inoculated from agar into YEPD complex media broth (100 ml) and grown overnight at 30°C. The YEPD broth was prepared as described in Kaiser *et al.*, *Methods in Yeast Genetics*, Cold Spring Harbor Press, Cold Spring Harbor, NY, p. 207 (1994). The overnight culture was then diluted to about 30 2 x 10⁶ cells/ml (approx. OD₆₀₀=0.1) into fresh YEPD broth (500 ml) and regrown to 1 x 10⁷ cells/ml (approx. OD₆₀₀=0.4-0.5).

30

The cells were then harvested and prepared for transformation by transfer into GS3 rotor bottles in a Sorval GS3 rotor at 5,000 rpm for 5 minutes, the supernatant discarded, and then resuspended into sterile water, and centrifuged again in 50 ml falcon tubes at 3,500 rpm in a Beckman GS-6KR centrifuge. The supernatant was discarded and the cells were subsequently washed with LiAc/TE (10 ml, 10 mM Tris-HCl, 1 mM EDTA pH 7.5, 100 mM Li₂OOCCH₃), and resuspended into LiAc/TE (2.5 ml).

35

Transformation took place by mixing the prepared cells (100 µl) with freshly denatured single stranded

salmon testes DNA (Lofstrand Labs, Gaithersburg, MD) and transforming DNA (1 μ g, vol. < 10 μ l) in microfuge tubes. The mixture was mixed briefly by vortexing, then 40% PEG/TE (600 μ l, 40% polyethylene glycol-4000, 10 mM Tris-HCl, 1 mM EDTA, 100 mM Li₂OOCCH₃, pH 7.5) was added. This mixture was gently mixed and incubated at 30°C while agitating for 30 minutes. The cells were then heat shocked at 42°C for 15 minutes, and the reaction vessel centrifuged in a microfuge at 12,000 rpm for 5-10 seconds, decanted and 5 resuspended into TE (500 μ l, 10 mM Tris-HCl, 1 mM EDTA pH 7.5) followed by recentrifugation. The cells were then diluted into TE (1 ml) and aliquots (200 μ l) were spread onto the selective media previously prepared in 150 mm growth plates (VWR).

Alternatively, instead of multiple small reactions, the transformation was performed using a single, large scale reaction, wherein reagent amounts were scaled up accordingly.

10 The selective media used was a synthetic complete dextrose agar lacking uracil (SCD-Ura) prepared as described in Kaiser *et al.*, Methods in Yeast Genetics, Cold Spring Harbor Press, Cold Spring Harbor, NY, p. 208-210 (1994). Transformants were grown at 30°C for 2-3 days.

15 The detection of colonies secreting amylase was performed by including red starch in the selective growth media. Starch was coupled to the red dye (Reactive Red-120, Sigma) as per the procedure described by Biely *et al.*, Anal. Biochem., 172:176-179 (1988). The coupled starch was incorporated into the SCD-Ura agar plates at a final concentration of 0.15% (w/v), and was buffered with potassium phosphate to a pH of 7.0 (50-100 mM final concentration).

20 The positive colonies were picked and streaked across fresh selective media (onto 150 mm plates) in order to obtain well isolated and identifiable single colonies. Well isolated single colonies positive for amylase secretion were detected by direct incorporation of red starch into buffered SCD-Ura agar. Positive colonies were determined by their ability to break down starch resulting in a clear halo around the positive colony visualized directly.

4. Isolation of DNA by PCR Amplification

25 When a positive colony was isolated, a portion of it was picked by a toothpick and diluted into sterile water (30 μ l) in a 96 well plate. At this time, the positive colonies were either frozen and stored for subsequent analysis or immediately amplified. An aliquot of cells (5 μ l) was used as a template for the PCR reaction in a 25 μ l volume containing: 0.5 μ l Klentaq (Clontech, Palo Alto, CA); 4.0 μ l 10 mM dNTP's (Perkin Elmer-Cetus); 2.5 μ l Kentaq buffer (Clontech); 0.25 μ l forward oligo 1; 0.25 μ l reverse oligo 2; 12.5 μ l distilled water. The 30 sequence of the forward oligonucleotide 1 was:

5'-TGTAAAACGACGGCCAGTTAAATAGACCTGCAATTATTAATCT-3' (SEQ ID NO:245)

The sequence of reverse oligonucleotide 2 was:

5'-CAGGAAACAGCTATGACCACCTGCACACCTGCAAATCCATT-3' (SEQ ID NO:246)

PCR was then performed as follows:

35	a.	Denature	92°C, 5 minutes
	b. 3 cycles of:	Denature	92°C, 30 seconds
		Anneal	59°C, 30 seconds

		Extend	72°C, 60 seconds
5	c. 3 cycles of:	Denature	92°C, 30 seconds
		Anneal	57°C, 30 seconds
		Extend	72°C, 60 seconds
10	d. 25 cycles of:	Denature	92°C, 30 seconds
		Anneal	55°C, 30 seconds
		Extend	72°C, 60 seconds
15	e.	Hold	4°C

The underlined regions of the oligonucleotides annealed to the ADH promoter region and the amylase region, respectively, and amplified a 307 bp region from vector pSST-AMY.0 when no insert was present. Typically, the first 18 nucleotides of the 5' end of these oligonucleotides contained annealing sites for the sequencing primers. Thus, the total product of the PCR reaction from an empty vector was 343 bp. However, signal sequence-fused cDNA resulted in considerably longer nucleotide sequences.

Following the PCR, an aliquot of the reaction (5 μ l) was examined by agarose gel electrophoresis in a 1% agarose gel using a Tris-Borate-EDTA (TBE) buffering system as described by Sambrook *et al.*, *supra*. Clones resulting in a single strong PCR product larger than 400 bp were further analyzed by DNA sequencing 20 after purification with a 96 Qiaquick PCR clean-up column (Qiagen Inc., Chatsworth, CA).

EXAMPLE 3: Isolation of cDNA Clones Using Signal Algorithm Analysis

Various polypeptide-encoding nucleic acid sequences were identified by applying a proprietary signal sequence finding algorithm developed by Genentech, Inc. (South San Francisco, CA) upon ESTs as well as 25 clustered and assembled EST fragments from public (*e.g.*, GenBank) and/or private (LIFESEQ[®], Incyte Pharmaceuticals, Inc., Palo Alto, CA) databases. The signal sequence algorithm computes a secretion signal score based on the character of the DNA nucleotides surrounding the first and optionally the second methionine codon(s) (ATG) at the 5'-end of the sequence or sequence fragment under consideration. The nucleotides following the first ATG must code for at least 35 unambiguous amino acids without any stop codons. If the first 30 ATG has the required amino acids, the second is not examined. If neither meets the requirement, the candidate sequence is not scored. In order to determine whether the EST sequence contains an authentic signal sequence, the DNA and corresponding amino acid sequences surrounding the ATG codon are scored using a set of seven sensors (evaluation parameters) known to be associated with secretion signals. Use of this algorithm resulted in the identification of numerous polypeptide-encoding nucleic acid sequences.

35

EXAMPLE 4: Isolation of cDNA clones Encoding Human PRO Polypeptides

Using the techniques described in Examples 1 to 3 above, numerous full-length cDNA clones were identified as encoding PRO polypeptides as disclosed herein. These cDNAs were then deposited under the terms of the Budapest Treaty with the American Type Culture Collection, 10801 University Blvd., Manassas, VA 40 20110-2209, USA (ATCC) as shown in Table 7 below.

Table 7

<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
DNA94849-2960	PTA-2306	July 25, 2000
DNA96883-2745	PTA-544	August 17, 1999
5 DNA96894-2675	PTA-260	June 22, 1999
DNA100272-2969	PTA-2299	July 25, 2000
DNA108696-2966	PTA-2315	August 1, 2000
DNA117935-2801	PTA-1088	December 22, 1999
DNA119474-2803	PTA-1097	December 22, 1999
10 DNA119498-2965	PTA-2298	July 25, 2000
DNA119502-2789	PTA-1082	December 22, 1999
DNA119516-2797	PTA-1083	December 22, 1999
DNA119530-2968	PTA-2396	August 8, 2000
DNA121772-2741	PTA-1030	December 7, 1999
15 DNA125148-2782	PTA-955	November 16, 1999
DNA125150-2793	PTA-1085	December 22, 1999
DNA125151-2784	PTA-1029	December 7, 1999
DNA125181-2804	PTA-1096	December 22, 1999
DNA125192-2794	PTA-1086	December 22, 1999
20 DNA125196-2792	PTA-1091	December 22, 1999
DNA125200-2810	PTA-1186	January 11, 2000
DNA125214-2814	PTA-1270	February 2, 2000
DNA125219-2799	PTA-1084	December 22, 1999
DNA128309-2825	PTA-1340	February 8, 2000
25 DNA129535-2796	PTA-1087	December 22, 1999
DNA129549-2798	PTA-1099	December 22, 1999
DNA129580-2863	PTA-1584	March 28, 2000
DNA129794-2967	PTA-2305	July 25, 2000
DNA131590-2962	PTA-2297	July 25, 2000
30 DNA135173-2811	PTA-1184	January 11, 2000
DNA138039-2828	PTA-1343	February 8, 2000
DNA139540-2807	PTA-1187	January 11, 2000
DNA139602-2859	PTA-1588	March 28, 2000
DNA139632-2880	PTA-1629	April 4, 2000
35 DNA139686-2823	PTA-1264	February 2, 2000
DNA142392-2800	PTA-1092	December 22, 1999
DNA143076-2787	PTA-1028	December 7, 1999

Table 7 (cont')

<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
DNA143294-2818	PTA-1182	January 11, 2000
DNA143514-2817	PTA-1266	February 2, 2000
DNA144841-2816	PTA-1188	January 11, 2000
5	DNA148380-2827	January 11, 2000
	DNA149995-2871	May 31, 2000
	DNA167678-2963	July 25, 2000
	DNA168028-2956	July 25, 2000
	DNA173894-2947	June 20, 2000
10	DNA176775-2957	July 25, 2000
	DNA177313-2982	July 19, 2000
	DNA57700-1408	January 12, 1999
	DNA62872-1509	August 4, 1998
	DNA62876-1517	August 4, 1998
15	DNA66660-1585	September 22, 1998
	DNA34434-1139	September 16, 1997
	DNA44804-1248	December 10, 1997
	DNA52758-1399	April 14, 1998
	DNA59849-1504	June 16, 1998
20	DNA65410-1569	September 15, 1998
	DNA71290-1630	September 22, 1998
	DNA33100-1159	October 16, 1997
	DNA64896-1539	September 9, 1998
	DNA84920-2614	April 27, 1999
25	DNA23330-1390	April 14, 1998
	DNA32286-1191	October 16, 1997
	DNA35673-1201	October 28, 1997
	DNA43316-1237	November 21, 1997
	DNA44184-1319	March 26, 1998
30	DNA45419-1252	February 5, 1998
	DNA48314-1320	March 26, 1998
	DNA50921-1458	May 12, 1998
	DNA53987	May 12, 1998
	DNA56047-1456	June 9, 1998
35	DNA56405-1357	May 6, 1998
	DNA56531-1648	September 29, 1998
	DNA56865-1491	June 23, 1998

Table 7 (cont')

	DNA57694-1341	203017	June 23, 1998
	DNA57708-1411	203021	June 23, 1998
	DNA57836-1338	203025	June 23, 1998
	DNA57841-1522	203458	November 3, 1998
5	DNA58847-1383	209879	May 20, 1998
	DNA59212-1627	203245	September 9, 1998
	DNA59588-1571	203106	August 11, 1998
	DNA59622-1334	209984	June 16, 1998
	DNA59847-2510	203576	January 12, 1999
	DNA60615-1483	209980	June 16, 1998
10	DNA60621-1516	203091	August 4, 1998
	DNA62814-1521	203093	August 4, 1998
	DNA64883-1526	203253	September 9, 1998
	DNA64889-1541	203250	September 9, 1998
	DNA64897-1628	203216	September 15, 1998
15	DNA64903-1553	203223	September 15, 1998
	DNA64907-1163-1	203242	September 9, 1998
	DNA64950-1590	203224	September 15, 1998
	DNA64952-1568	203222	September 15, 1998
	DNA65402-1540	203252	September 9, 1998
20	DNA65405-1547	203476	November 17, 1998
	DNA66663-1598	203268	September 22, 1998
	DNA66667	203267	September 22, 1998
	DNA66675-1587	203282	September 22, 1998
	DNA67300-1605	203163	August 25, 1998
25	DNA68872-1620	203160	August 25, 1998
	DNA71269-1621	203284	September 22, 1998
	DNA73736-1657	203466	November 17, 1998
	DNA73739-1645	203270	September 22, 1998
	DNA76400-2528	203573	January 12, 1999
30	DNA76532-1702	203473	November 17, 1998
	DNA76541-1675	203409	October 27, 1998
	DNA79862-2522	203550	December 22, 1998
	DNA81754-2532	203542	December 15, 1998
	DNA81761-2583	203862	March 23, 1999
35	DNA83500-2506	203391	October 29, 1998
	DNA84210-2576	203818	March 2, 1999

Table 7 (cont')

DNA86571-2551	203660	February 9, 1999
DNA92218-2554	203834	March 9, 1999
DNA92223-2567	203851	March 16, 1999
DNA92265-2669	PTA-256	June 22, 1999
5	DNA92274-2617	April 27, 1999
	DNA108760-2740	August 17, 1999
	DNA108792-2753	August 31, 1999
	DNA111750-2706	August 3, 1999
	DNA119514-2772	November 9, 1999
	DNA125185-2806	December 7, 1999

These deposits were made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the deposit for 30 years from the date of deposit. The 15 deposits will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures permanent and unrestricted availability of the progeny of the culture of the deposit to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent application, whichever comes first, and assures availability of the progeny to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 USC § 20 122 and the Commissioner's rules pursuant thereto (including 37 CFR § 1.14 with particular reference to 886 OG 638).

The assignee of the present application has agreed that if a culture of the materials on deposit should die or be lost or destroyed when cultivated under suitable conditions, the materials will be promptly replaced on notification with another of the same. Availability of the deposited material is not to be construed as a license 25 to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

EXAMPLE 5: Isolation of cDNA clones Encoding Human PRO6004, PRO5723, PRO3444, and PRO9940

DNA molecules encoding the PRO840, PRO1338, PRO6004, PRO5723, PRO3444, and PRO9940 30 polypeptides shown in the accompanying figures were obtained through GenBank.

EXAMPLE 6: Use of PRO as a hybridization probe

The following method describes use of a nucleotide sequence encoding PRO as a hybridization probe.

DNA comprising the coding sequence of full-length or mature PRO as disclosed herein is employed as 35 a probe to screen for homologous DNAs (such as those encoding naturally-occurring variants of PRO) in human tissue cDNA libraries or human tissue genomic libraries.

Hybridization and washing of filters containing either library DNAs is performed under the following

high stringency conditions. Hybridization of radiolabeled PRO-derived probe to the filters is performed in a solution of 50% formamide, 5x SSC, 0.1% SDS, 0.1% sodium pyrophosphate, 50 mM sodium phosphate, pH 6.8, 2x Denhardt's solution, and 10% dextran sulfate at 42°C for 20 hours. Washing of the filters is performed in an aqueous solution of 0.1x SSC and 0.1% SDS at 42°C.

5 DNAs having a desired sequence identity with the DNA encoding full-length native sequence PRO can then be identified using standard techniques known in the art.

EXAMPLE 7: Expression of PRO in *E. coli*

This example illustrates preparation of an unglycosylated form of PRO by recombinant expression in *E. coli*.

10 The DNA sequence encoding PRO is initially amplified using selected PCR primers. The primers should contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is pBR322 (derived from *E. coli*; see Bolivar et al., *Gene*, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated 15 into the vector. The vector will preferably include sequences which encode for an antibiotic resistance gene, a trp promoter, a polyhis leader (including the first six STII codons, polyhis sequence, and enterokinase cleavage site), the PRO coding region, lambda transcriptional terminator, and an argU gene.

20 The ligation mixture is then used to transform a selected *E. coli* strain using the methods described in Sambrook et al., *supra*. Transformants are identified by their ability to grow on LB plates and antibiotic resistant colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

25 After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized PRO protein can then be purified using a metal chelating column under conditions that allow tight binding of the protein.

30 PRO may be expressed in *E. coli* in a poly-His tagged form, using the following procedure. The DNA encoding PRO is initially amplified using selected PCR primers. The primers will contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector, and other useful sequences providing for efficient and reliable translation initiation, rapid purification on a metal chelation column, and proteolytic removal with enterokinase. The PCR-amplified, poly-His tagged sequences are then ligated into an expression vector, which is used to transform an *E. coli* host based on strain 52 (W3110 fuhA(tonA) lon galE 35 rpoHts(htpRts) clpP(lacIq)). Transformants are first grown in LB containing 50 mg/ml carbenicillin at 30°C with shaking until an O.D.600 of 3-5 is reached. Cultures are then diluted 50-100 fold into CRAP media (prepared by mixing 3.57 g (NH₄)₂SO₄, 0.71 g sodium citrate•2H₂O, 1.07 g KCl, 5.36 g Difco yeast extract, 5.36 g

Sheffield hycase SF in 500 mL water, as well as 110 mM MPOS, pH 7.3, 0.55% (w/v) glucose and 7 mM MgSO₄) and grown for approximately 20-30 hours at 30°C with shaking. Samples are removed to verify expression by SDS-PAGE analysis, and the bulk culture is centrifuged to pellet the cells. Cell pellets are frozen until purification and refolding.

5 *E. coli* paste from 0.5 to 1 L fermentations (6-10 g pellets) is resuspended in 10 volumes (w/v) in 7 M guanidine, 20 mM Tris, pH 8 buffer. Solid sodium sulfite and sodium tetrathionate is added to make final concentrations of 0.1M and 0.02 M, respectively, and the solution is stirred overnight at 4°C. This step results in a denatured protein with all cysteine residues blocked by sulfitolization. The solution is centrifuged at 40,000 rpm in a Beckman Ultracentrifuge for 30 min. The supernatant is diluted with 3-5 volumes of metal chelate column buffer (6 M guanidine, 20 mM Tris, pH 7.4) and filtered through 0.22 micron filters to clarify. The 10 clarified extract is loaded onto a 5 ml Qiagen Ni-NTA metal chelate column equilibrated in the metal chelate column buffer. The column is washed with additional buffer containing 50 mM imidazole (Calbiochem, Utrol grade), pH 7.4. The protein is eluted with buffer containing 250 mM imidazole. Fractions containing the desired protein are pooled and stored at 4°C. Protein concentration is estimated by its absorbance at 280 nm using the calculated extinction coefficient based on its amino acid sequence.

15 The proteins are refolded by diluting the sample slowly into freshly prepared refolding buffer consisting of: 20 mM Tris, pH 8.6, 0.3 M NaCl, 2.5 M urea, 5 mM cysteine, 20 mM glycine and 1 mM EDTA. Refolding volumes are chosen so that the final protein concentration is between 50 to 100 micrograms/ml. The refolding solution is stirred gently at 4°C for 12-36 hours. The refolding reaction is quenched by the addition of TFA to a final concentration of 0.4% (pH of approximately 3). Before further purification of the protein, the solution 20 is filtered through a 0.22 micron filter and acetonitrile is added to 2-10% final concentration. The refolded protein is chromatographed on a Poros R1/H reversed phase column using a mobile buffer of 0.1% TFA with elution with a gradient of acetonitrile from 10 to 80%. Aliquots of fractions with A280 absorbance are analyzed on SDS polyacrylamide gels and fractions containing homogeneous refolded protein are pooled. Generally, the 25 properly refolded species of most proteins are eluted at the lowest concentrations of acetonitrile since those species are the most compact with their hydrophobic interiors shielded from interaction with the reversed phase resin. Aggregated species are usually eluted at higher acetonitrile concentrations. In addition to resolving misfolded forms of proteins from the desired form, the reversed phase step also removes endotoxin from the samples.

30 Fractions containing the desired folded PRO polypeptide are pooled and the acetonitrile removed using a gentle stream of nitrogen directed at the solution. Proteins are formulated into 20 mM Hepes, pH 6.8 with 0.14 M sodium chloride and 4% mannitol by dialysis or by gel filtration using G25 Superfine (Pharmacia) resins equilibrated in the formulation buffer and sterile filtered.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

EXAMPLE 8: Expression of PRO in mammalian cells

35 This example illustrates preparation of a potentially glycosylated form of PRO by recombinant expression in mammalian cells.

The vector, pRK5 (see EP 307,247, published March 15, 1989), is employed as the expression vector.

Optionally, the PRO DNA is ligated into pRK5 with selected restriction enzymes to allow insertion of the PRO DNA using ligation methods such as described in Sambrook et al., *supra*. The resulting vector is called pRK5-PRO.

In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and 5 optional, nutrient components and/or antibiotics. About 10 μ g pRK5-PRO DNA is mixed with about 1 μ g DNA encoding the VA RNA gene [Thimmappaya et al., *Cell*, 31:543 (1982)] and dissolved in 500 μ l of 1 mM Tris-HCl, 0.1 mM EDTA, 0.227 M CaCl₂. To this mixture is added, dropwise, 500 μ l of 50 mM HEPES (pH 7.35), 10 280 mM NaCl, 1.5 mM NaPO₄, and a precipitate is allowed to form for 10 minutes at 25°C. The precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37°C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with serum free medium, fresh medium is added and the cells are incubated for about 5 days.

15 Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture medium (alone) or culture medium containing 200 μ Ci/ml ³⁵S-cysteine and 200 μ Ci/ml ³⁵S-methionine. After a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel may be dried and exposed to film for a selected period of time to reveal the presence 20 of PRO polypeptide. The cultures containing transfected cells may undergo further incubation (in serum free medium) and the medium is tested in selected bioassays.

25 In an alternative technique, PRO may be introduced into 293 cells transiently using the dextran sulfate method described by Sompanyrac et al., *Proc. Natl. Acad. Sci.*, 78:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700 μ g pRK5-PRO DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5 μ g/ml bovine insulin and 0.1 μ g/ml bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris. The sample containing expressed PRO can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography.

30 In another embodiment, PRO can be expressed in CHO cells. The pRK5-PRO can be transfected into CHO cells using known reagents such as CaPO₄ or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as ³⁵S-methionine. After determining the presence of PRO polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. The medium containing the expressed PRO can then be concentrated and purified by any selected method.

35 Epitope-tagged PRO may also be expressed in host CHO cells. The PRO may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-his tag into a Baculovirus expression vector. The poly-his tagged PRO insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells

can be transfected (as described above) with the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO can then be concentrated and purified by any selected method, such as by Ni^{2+} -chelate affinity chromatography.

PRO may also be expressed in CHO and/or COS cells by a transient expression procedure or in CHO cells by another stable expression procedure.

5 Stable expression in CHO cells is performed using the following procedure. The proteins are expressed as an IgG construct (immunoadhesin), in which the coding sequences for the soluble forms (e.g. extracellular domains) of the respective proteins are fused to an IgG1 constant region sequence containing the hinge, CH2 and CH2 domains and/or is a poly-His tagged form.

Following PCR amplification, the respective DNAs are subcloned in a CHO expression vector using 10 standard techniques as described in Ausubel et al., *Current Protocols of Molecular Biology*, Unit 3.16, John Wiley and Sons (1997). CHO expression vectors are constructed to have compatible restriction sites 5' and 3' of the DNA of interest to allow the convenient shuttling of cDNA's. The vector used expression in CHO cells is as described in Lucas et al., *Nucl. Acids Res.* 24:9 (1774-1779 (1996), and uses the SV40 early promoter/enhancer to drive expression of the cDNA of interest and dihydrofolate reductase (DHFR). DHFR expression permits 15 selection for stable maintenance of the plasmid following transfection.

Twelve micrograms of the desired plasmid DNA is introduced into approximately 10 million CHO cells using commercially available transfection reagents Superfect[®] (Qiagen), Dasurer[®] or Fugene[®] (Boehringer Mannheim). The cells are grown as described in Lucas et al., *supra*. Approximately 3×10^7 cells are frozen in an ampule for further growth and production as described below.

20 The ampules containing the plasmid DNA are thawed by placement into water bath and mixed by vortexing. The contents are pipetted into a centrifuge tube containing 10 mLs of media and centrifuged at 1000 rpm for 5 minutes. The supernatant is aspirated and the cells are resuspended in 10 mL of selective media (0.2 μm filtered PS20 with 5% 0.2 μm diafiltered fetal bovine serum). The cells are then aliquoted into a 100 mL spinner containing 90 mL of selective media. After 1-2 days, the cells are transferred into a 250 mL spinner filled 25 with 150 mL selective growth medium and incubated at 37°C. After another 2-3 days, 250 mL, 500 mL and 2000 mL spinners are seeded with 3×10^5 cells/mL. The cell media is exchanged with fresh media by centrifugation and resuspension in production medium. Although any suitable CHO media may be employed, a production medium described in U.S. Patent No. 5,122,469, issued June 16, 1992 may actually be used. A 3L production spinner is seeded at 1.2×10^6 cells/mL. On day 0, the cell number pH is determined. On day 1, the spinner is 30 sampled and sparging with filtered air is commenced. On day 2, the spinner is sampled, the temperature shifted to 33°C, and 30 mL of 500 g/L glucose and 0.6 mL of 10% antifoam (e.g., 35% polydimethylsiloxane emulsion, Dow Corning 365 Medical Grade Emulsion) taken. Throughout the production, the pH is adjusted as necessary to keep it at around 7.2. After 10 days, or until the viability dropped below 70%, the cell culture is harvested by centrifugation and filtering through a 0.22 μm filter. The filtrate was either stored at 4°C or immediately 35 loaded onto columns for purification.

For the poly-His tagged constructs, the proteins are purified using a Ni-NTA column (Qiagen). Before purification, imidazole is added to the conditioned media to a concentration of 5 mM. The conditioned media is

pumped onto a 6 ml Ni-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 ml/min. at 4°C. After loading, the column is washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly purified protein is subsequently desalted into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 ml G25 Superfine (Pharmacia) column and stored at -80°C.

5 Immunoadhesin (Fc-containing) constructs are purified from the conditioned media as follows. The conditioned medium is pumped onto a 5 ml Protein A column (Pharmacia) which had been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column is washed extensively with equilibration buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein is immediately neutralized by collecting 1 ml fractions into tubes containing 275 μ L of 1 M Tris buffer, pH 9. The highly purified protein is subsequently 10 desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity is assessed by SDS polyacrylamide gels and by N-terminal amino acid sequencing by Edman degradation.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

EXAMPLE 9: Expression of PRO in Yeast

15 The following method describes recombinant expression of PRO in yeast.

First, yeast expression vectors are constructed for intracellular production or secretion of PRO from the ADH2/GAPDH promoter. DNA encoding PRO and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of PRO. For secretion, DNA encoding PRO can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, a native PRO signal 20 peptide or other mammalian signal peptide, or, for example, a yeast alpha-factor or invertase secretory signal/leader sequence, and linker sequences (if needed) for expression of PRO.

Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with 25 Coomassie Blue stain.

Recombinant PRO can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing PRO may further be purified using selected column chromatography resins.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

30

EXAMPLE 10: Expression of PRO in Baculovirus-Infected Insect Cells

The following method describes recombinant expression of PRO in Baculovirus-infected insect cells.

The sequence coding for PRO is fused upstream of an epitope tag contained within a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG). 35 A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the sequence encoding PRO or the desired portion of the coding sequence of PRO such as the sequence encoding the extracellular domain of a transmembrane protein or the sequence encoding

the mature protein if the protein is extracellular is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

5 Recombinant baculovirus is generated by co-transfected the above plasmid and BaculoGold™ virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4 - 5 days of incubation at 28°C, the released viruses are harvested and used for further amplifications. Viral infection and protein expression are performed as described by O'Reilley et al., Baculovirus expression vectors: A Laboratory Manual, Oxford: Oxford University Press (1994).

10 Expressed poly-his tagged PRO can then be purified, for example, by Ni²⁺-chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert et al., Nature, 362:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 mL Hepes, pH 7.9; 12.5 mM MgCl₂; 0.1 mM EDTA; 10% glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated twice for 20 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% glycerol, pH 7.8) and filtered through a 0.45 μm filter. A Ni²⁺-NTA agarose column (commercially available from Qiagen) is prepared with a bed volume of 5 mL, 15 washed with 25 mL of water and equilibrated with 25 mL of loading buffer. The filtered cell extract is loaded onto the column at 0.5 mL per minute. The column is washed to baseline A₂₈₀ with loading buffer, at which point fraction collection is started. Next, the column is washed with a secondary wash buffer (50 mM phosphate; 300 mM NaCl, 10% glycerol, pH 6.0), which elutes nonspecifically bound protein. After reaching A₂₈₀ baseline again, the column is developed with a 0 to 500 mM Imidazole gradient in the secondary wash buffer. One mL 20 fractions are collected and analyzed by SDS-PAGE and silver staining or Western blot with Ni²⁺-NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His₁₀-tagged PRO are pooled and dialyzed against loading buffer.

Alternatively, purification of the IgG tagged (or Fc tagged) PRO can be performed using known chromatography techniques, including for instance, Protein A or protein G column chromatography.

25 Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

EXAMPLE 11: Preparation of Antibodies that Bind PRO

This example illustrates preparation of monoclonal antibodies which can specifically bind PRO.

30 Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, supra. Immunogens that may be employed include purified PRO, fusion proteins containing PRO, and cells expressing recombinant PRO on the cell surface. Selection of the immunogen can be made by the skilled artisan without undue experimentation.

35 Mice, such as Balb/c, are immunized with the PRO immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with

additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect anti-PRO antibodies.

5 After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of PRO. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen cells are then fused (using 35% polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then be plated in 96 well tissue culture plates containing HAT (hypoxanthine, aminopterin, and thymidine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

The hybridoma cells will be screened in an ELISA for reactivity against PRO. Determination of "positive" hybridoma cells secreting the desired monoclonal antibodies against PRO is within the skill in the art.

10 The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-PRO monoclonal antibodies. Alternatively, the hybridoma cells can be grown in tissue culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be accomplished using ammonium sulfate precipitation, followed by gel exclusion chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

15

EXAMPLE 12: Purification of PRO Polypeptides Using Specific Antibodies

Native or recombinant PRO polypeptides may be purified by a variety of standard techniques in the art of protein purification. For example, pro-PRO polypeptide, mature PRO polypeptide, or pre-PRO polypeptide is purified by immunoaffinity chromatography using antibodies specific for the PRO polypeptide of interest. In 20 general, an immunoaffinity column is constructed by covalently coupling the anti-PRO polypeptide antibody to an activated chromatographic resin.

25 Polyclonal immunoglobulins are prepared from immune sera either by precipitation with ammonium sulfate or by purification on immobilized Protein A (Pharmacia LKB Biotechnology, Piscataway, N.J.). Likewise, monoclonal antibodies are prepared from mouse ascites fluid by ammonium sulfate precipitation or chromatography on immobilized Protein A. Partially purified immunoglobulin is covalently attached to a chromatographic resin such as CnBr-activated SEPHAROSETM (Pharmacia LKB Biotechnology). The antibody is coupled to the resin, the resin is blocked, and the derivative resin is washed according to the manufacturer's instructions.

Such an immunoaffinity column is utilized in the purification of PRO polypeptide by preparing a fraction 30 from cells containing PRO polypeptide in a soluble form. This preparation is derived by solubilization of the whole cell or of a subcellular fraction obtained via differential centrifugation by the addition of detergent or by other methods well known in the art. Alternatively, soluble PRO polypeptide containing a signal sequence may be secreted in useful quantity into the medium in which the cells are grown.

35 A soluble PRO polypeptide-containing preparation is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of PRO polypeptide (e.g., high ionic strength buffers in the presence of detergent). Then, the column is eluted under conditions that disrupt antibody/PRO polypeptide binding (e.g., a low pH buffer such as approximately pH 2-3, or a high concentration

of a chaotrope such as urea or thiocyanate ion), and PRO polypeptide is collected.

EXAMPLE 13: Drug Screening

This invention is particularly useful for screening compounds by using PRO polypeptides or binding fragment thereof in any of a variety of drug screening techniques. The PRO polypeptide or fragment employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the PRO polypeptide or fragment. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between PRO polypeptide or a fragment and the agent being tested. Alternatively, one can examine the diminution in complex formation between the PRO polypeptide and its target cell or target receptors caused by the agent being tested.

Thus, the present invention provides methods of screening for drugs or any other agents which can affect a PRO polypeptide-associated disease or disorder. These methods comprise contacting such an agent with an PRO polypeptide or fragment thereof and assaying (I) for the presence of a complex between the agent and the PRO polypeptide or fragment, or (ii) for the presence of a complex between the PRO polypeptide or fragment and the cell, by methods well known in the art. In such competitive binding assays, the PRO polypeptide or fragment is typically labeled. After suitable incubation, free PRO polypeptide or fragment is separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the ability of the particular agent to bind to PRO polypeptide or to interfere with the PRO polypeptide/cell complex.

Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity to a polypeptide and is described in detail in WO 84/03564, published on September 13, 1984. Briefly stated, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. As applied to a PRO polypeptide, the peptide test compounds are reacted with PRO polypeptide and washed. Bound PRO polypeptide is detected by methods well known in the art. Purified PRO polypeptide can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the peptide and immobilize it on the solid support.

This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding PRO polypeptide specifically compete with a test compound for binding to PRO polypeptide or fragments thereof. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with PRO polypeptide.

EXAMPLE 14: Rational Drug Design

The goal of rational drug design is to produce structural analogs of biologically active polypeptide of interest (*i.e.*, a PRO polypeptide) or of small molecules with which they interact, *e.g.*, agonists, antagonists, or inhibitors. Any of these examples can be used to fashion drugs which are more active or stable forms of the PRO polypeptide or which enhance or interfere with the function of the PRO polypeptide *in vivo* (*c.f.*, Hodgson,

Bio/Technology, 9: 19-21 (1991)).

In one approach, the three-dimensional structure of the PRO polypeptide, or of an PRO polypeptide-inhibitor complex, is determined by x-ray crystallography, by computer modeling or, most typically, by a combination of the two approaches. Both the shape and charges of the PRO polypeptide must be ascertained to elucidate the structure and to determine active site(s) of the molecule. Less often, useful information regarding the structure of the PRO polypeptide may be gained by modeling based on the structure of homologous proteins. In both cases, relevant structural information is used to design analogous PRO polypeptide-like molecules or to identify efficient inhibitors. Useful examples of rational drug design may include molecules which have improved activity or stability as shown by Braxton and Wells, Biochemistry, 31:7796-7801 (1992) or which act as inhibitors, agonists, or antagonists of native peptides as shown by Athauda *et al.*, J. Biochem., 113:742-746 (1993).

It is also possible to isolate a target-specific antibody, selected by functional assay, as described above, and then to solve its crystal structure. This approach, in principle, yields a pharmacore upon which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original receptor. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced peptides. The isolated peptides would then act as the pharmacore.

By virtue of the present invention, sufficient amounts of the PRO polypeptide may be made available to perform such analytical studies as X-ray crystallography. In addition, knowledge of the PRO polypeptide amino acid sequence provided herein will provide guidance to those employing computer modeling techniques in place of or in addition to x-ray crystallography.

EXAMPLE 15: Pericyte c-Fos Induction (Assay 93)

This assay shows that certain polypeptides of the invention act to induce the expression of c-fos in pericyte cells and, therefore, are useful not only as diagnostic markers for particular types of pericyte-associated tumors but also for giving rise to antagonists which would be expected to be useful for the therapeutic treatment of pericyte-associated tumors. Induction of c-fos expression in pericytes is also indicative of the induction of angiogenesis and, as such, PRO polypeptides capable of inducing the expression of c-fos would be expected to be useful for the treatment of conditions where induced angiogenesis would be beneficial including, for example, wound healing, and the like. Specifically, on day 1, pericytes are received from VEC Technologies and all but 5 ml of media is removed from flask. On day 2, the pericytes are trypsinized, washed, spun and then plated onto 96 well plates. On day 7, the media is removed and the pericytes are treated with 100 μ l of PRO polypeptide test samples and controls (positive control = DME+5% serum +/- PDGF at 500 ng/ml; negative control = protein 32). Replicates are averaged and SD/CV are determined. Fold increase over Protein 32 (buffer control) value indicated by chemiluminescence units (RLU) luminometer reading versus frequency is plotted on a histogram. Two-fold above Protein 32 value is considered positive for the assay. ASY Matrix: Growth media = low glucose DMEM = 20% FBS + 1X pen strep + 1X fungizone. Assay Media = low glucose DMEM + 5% FBS.

The following polypeptides tested positive in this assay: PRO982, PRO1160, PRO1187, and PRO1329.

EXAMPLE 16: Chondrocyte Re-differentiation Assay (Assay 110)

This assay shows that certain polypeptides of the invention act to induce redifferentiation of chondrocytes, therefore, are expected to be useful for the treatment of various bone and/or cartilage disorders such as, for example, sports injuries and arthritis. The assay is performed as follows. Porcine chondrocytes are isolated by overnight collagenase digestion of articular cartilage of metacarpophalangeal joints of 4-6 month old female pigs.

5 The isolated cells are then seeded at 25,000 cells/cm² in Ham F-12 containing 10% FBS and 4 µg/ml gentamycin. The culture media is changed every third day and the cells are then seeded in 96 well plates at 5,000 cells/well in 100µl of the same media without serum and 100 µl of the test PRO polypeptide, 5 nM staurosporin (positive control) or medium alone (negative control) is added to give a final volume of 200 µl/well. After 5 days of incubation at 37°C, a picture of each well is taken and the differentiation state of the chondrocytes is determined.

10 10 A positive result in the assay occurs when the redifferentiation of the chondrocytes is determined to be more similar to the positive control than the negative control.

The following polypeptide tested positive in this assay: PRO357.

EXAMPLE 17: Identification of PRO Polypeptides That Stimulate TNF-α Release In Human Blood (Assay 128)

15 This assay shows that certain PRO polypeptides of the present invention act to stimulate the release of TNF-α in human blood. PRO polypeptides testing positive in this assay are useful for, among other things, research purposes where stimulation of the release of TNF-α would be desired and for the therapeutic treatment of conditions wherein enhanced TNF-α release would be beneficial. Specifically, 200 µl of human blood supplemented with 50mM Hepes buffer (pH 7.2) is aliquoted per well in a 96 well test plate. To each well is then 20 added 300µl of either the test PRO polypeptide in 50 mM Hepes buffer (at various concentrations) or 50 mM Hepes buffer alone (negative control) and the plates are incubated at 37°C for 6 hours. The samples are then centrifuged and 50µl of plasma is collected from each well and tested for the presence of TNF-α by ELISA assay. A positive in the assay is a higher amount of TNF-α in the PRO polypeptide treated samples as compared to the negative control samples.

25 The following PRO polypeptides tested positive in this assay: PRO231, PRO357, PRO725, PRO1155, PRO1306, and PRO1419.

EXAMPLE 18: Promotion of Chondrocyte Redifferentiation (Assay 129)

30 This assay is designed to determine whether PRO polypeptides of the present invention show the ability to induce the proliferation and/or redifferentiation of chondrocytes in culture. PRO polypeptides testing positive in this assay would be expected to be useful for the therapeutic treatment of various bone and/or cartilage disorders such as, for example, sports injuries and arthritis.

35 Porcine chondrocytes are isolated by overnight collagenase digestion of articular cartilage of the metacarpophalangeal joint of 4-6 month old female pigs. The isolated cells are then seeded at 25,000 cells/cm² in Ham F-12 containing 10% FBS and 4 µg/ml gentamycin. The culture media is changed every third day. On day 12, the cells are seeded in 96 well plates at 5,000 cells/well in 100µl of the same media without serum and 100 µl of either serum-free medium (negative control), staurosporin (final concentration of 5 nM; positive control)

or the test PRO polypeptide are added to give a final volume of 200 μ l/well. After 5 days at 37°C, 22 μ l of media containing 100 μ g/ml Hoechst 33342 and 50 μ g/ml 5-CFDA is added to each well and incubated for an additional 10 minutes at 37°C. A picture of the green fluorescence is taken for each well and the differentiation state of the chondrocytes is calculated by morphometric analysis. A positive result in the assay is obtained when the > 50% of the PRO polypeptide treated cells are differentiated (compared to the background obtained by the negative control).

5 The following PRO polypeptides tested positive in this assay: PRO229, PRO1272, and PRO4405.

EXAMPLE 19: Normal Human Dermal Fibroblast Proliferation (Assay 141)

10 This assay is designed to determine whether PRO polypeptides of the present invention show the ability to induce proliferation of human dermal fibroblast cells in culture and, therefore, function as useful growth factors.

15 On day 0, human dermal fibroblast cells (from cell lines, maximum of 12-14 passages) were plated in 96-well plates at 1000 cells/well per 100 microliter and incubated overnight in complete media [fibroblast growth media (FGM, Clonetics), plus supplements: insulin, human epithelial growth factor (hEGF), gentamicin (GA-1000), and fetal bovine serum (FBS, Clonetics)]. On day 1, complete media was replaced by basal media [FGM plus 1% FBS] and addition of PRO polypeptides at 1%, 0.1% and 0.01%. On day 7, an assessment of cell proliferation was performed by Alamar Blue assay followed by Crystal Violet. Results are expresses as % of the cell growth observed with control buffer.

20 The following PRO polypeptides tested positive in this assay: PRO982, PRO357, PRO725, PRO1306, PRO1419, PRO229, PRO1272, PRO181, PRO214, PRO247, PRO337, PRO526, PRO363, PRO531, PRO1083, PRO840, PRO1080, PRO788, PRO1478, PRO1134, PRO826, PRO1005, PRO809, PRO1194, PRO1071, PRO1411, PRO1309, PRO1025, PRO1181, PRO1126, PRO1186, PRO1192, PRO1244, PRO1274, PRO1412, PRO1286, PRO1330, PRO1347, PRO1305, PRO1273, PRO1279, PRO1340, PRO1338, PRO1343, PRO1376, PRO1387, PRO1409, PRO1488, PRO1474, PRO1917, PRO1760, PRO1567, PRO1887, PRO1928, PRO4341, PRO5723, PRO1801, PRO4333, PRO3543, PRO3444, PRO4302, PRO4322, PRO5725, PRO4408, PRO9940, PRO7154, PRO7425, PRO6079, PRO9836 and PRO10096.

EXAMPLE 20: Microarray Analysis to Detect Overexpression of PRO Polypeptides in Cancerous Tumors

30 Nucleic acid microarrays, often containing thousands of gene sequences, are useful for identifying differentially expressed genes in diseased tissues as compared to their normal counterparts. Using nucleic acid microarrays, test and control mRNA samples from test and control tissue samples are reverse transcribed and labeled to generate cDNA probes. The cDNA probes are then hybridized to an array of nucleic acids immobilized on a solid support. The array is configured such that the sequence and position of each member of the array is known. For example, a selection of genes known to be expressed in certain disease states may be arrayed on a 35 solid support. Hybridization of a labeled probe with a particular array member indicates that the sample from which the probe was derived expresses that gene. If the hybridization signal of a probe from a test (disease tissue) sample is greater than hybridization signal of a probe from a control (normal tissue) sample, the gene or genes

overexpressed in the disease tissue are identified. The implication of this result is that an overexpressed protein in a diseased tissue is useful not only as a diagnostic marker for the presence of the disease condition, but also as a therapeutic target for treatment of the disease condition.

The methodology of hybridization of nucleic acids and microarray technology is well known in the art. In the present example, the specific preparation of nucleic acids for hybridization and probes, slides, and hybridization conditions are all detailed in U.S. Provisional Patent Application Serial No. 60/193,767, filed on March 31, 2000 and which is herein incorporated by reference.

In the present example, cancerous tumors derived from various human tissues were studied for PRO polypeptide-encoding gene expression relative to non-cancerous human tissue in an attempt to identify those PRO polypeptides which are overexpressed in cancerous tumors. Cancerous human tumor tissue from any of a variety 10 of different human tumors was obtained and compared to a "universal" epithelial control sample which was prepared by pooling non-cancerous human tissues of epithelial origin, including liver, kidney, and lung. mRNA isolated from the pooled tissues represents a mixture of expressed gene products from these different tissues. Microarray hybridization experiments using the pooled control samples generated a linear plot in a 2-color analysis. The slope of the line generated in a 2-color analysis was then used to normalize the ratios of (test:control 15 detection) within each experiment. The normalized ratios from various experiments were then compared and used to identify clustering of gene expression. Thus, the pooled "universal control" sample not only allowed effective relative gene expression determinations in a simple 2-sample comparison, it also allowed multi-sample comparisons across several experiments.

In the present experiments, nucleic acid probes derived from the herein described PRO polypeptide-encoding nucleic acid sequences were used in the creation of the microarray and RNA from a panel of nine 20 different tumor tissues (listed below) were used for the hybridization thereto. A value based upon the normalized ratio:experimental ratio was designated as a "cutoff ratio". Only values that were above this cutoff ratio were determined to be significant. Table 8 below shows the results of these experiments, demonstrating that various 25 PRO polypeptides of the present invention are significantly overexpressed in various human tumor tissues, as compared to a non-cancerous human tissue control or other human tumor tissues. As described above, these data demonstrate that the PRO polypeptides of the present invention are useful not only as diagnostic markers for the presence of one or more cancerous tumors, but also serve as therapeutic targets for the treatment of those tumors.

TABLE 8

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to normal control:</u>
30	PRO6004	colon tumor	universal normal control
	PRO4981	colon tumor	universal normal control
	PRO4981	lung tumor	universal normal control
35	PRO7174	colon tumor	universal normal control
	PRO5778	lung tumor	universal normal control
	PRO5778	breast tumor	universal normal control
40	PRO5778	liver tumor	universal normal control

TABLE 8 (cont')

<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to normal control:</u>
PRO4332	colon tumor	universal normal control
PRO9799	colon tumor	universal normal control
5 PRO9909	colon tumor	universal normal control
PRO9917	colon tumor	universal normal control
10 PRO9917	lung tumor	universal normal control
PRO9917	breast tumor	universal normal control
PRO9771	colon tumor	universal normal control
15 PRO9877	colon tumor	universal normal control
PRO9903	colon tumor	universal normal control
PRO9830	colon tumor	universal normal control
20 PRO7155	colon tumor	universal normal control
PRO7155	lung tumor	universal normal control
PRO7155	prostate tumor	universal normal control
25 PRO9862	colon tumor	universal normal control
PRO9882	colon tumor	universal normal control
PRO9864	colon tumor	universal normal control
30 PRO10013	colon tumor	universal normal control
PRO9885	colon tumor	universal normal control
PRO9879	colon tumor	universal normal control
35 PRO10111	colon tumor	universal normal control
PRO10111	rectal tumor	universal normal control
40 PRO9925	breast tumor	universal normal control
PRO9925	rectal tumor	universal normal control
PRO9925	colon tumor	universal normal control
PRO9925	lung tumor	universal normal control
45 PRO9905	colon tumor	universal normal control
PRO10276	colon tumor	universal normal control
PRO9898	colon tumor	universal normal control
50 PRO9904	colon tumor	universal normal control
PRO19632	colon tumor	universal normal control
PRO19672	colon tumor	universal normal control
55		

TABLE 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to normal control:</u>
5	PRO9783	colon tumor	universal normal control
	PRO9783	lung tumor	universal normal control
	PRO9783	breast tumor	universal normal control
	PRO9783	prostate tumor	universal normal control
	PRO9783	rectal tumor	universal normal control
	PRO10112	colon tumor	universal normal control
10	PRO10284	colon tumor	universal normal control
	PRO10100	colon tumor	universal normal control
15	PRO19628	colon tumor	universal normal control
	PRO19684	colon tumor	universal normal control
	PRO10274	colon tumor	universal normal control
	PRO9907	colon tumor	universal normal control
	PRO9873	colon tumor	universal normal control
25	PRO10201	colon tumor	universal normal control
	PRO10200	colon tumor	universal normal control
	PRO10196	colon tumor	universal normal control
30	PRO10282	lung tumor	universal normal control
	PRO10282	breast tumor	universal normal control
	PRO10282	colon tumor	universal normal control
	PRO10282	rectal tumor	universal normal control
35	PRO19650	colon tumor	universal normal control
40	PRO21184	lung tumor	universal normal control
	PRO21184	breast tumor	universal normal control
	PRO21184	colon tumor	universal normal control
	PRO21201	breast tumor	universal normal control
	PRO21201	colon tumor	universal normal control
45	PRO21175	breast tumor	universal normal control
	PRO21175	colon tumor	universal normal control
	PRO21175	lung tumor	universal normal control
50	PRO21340	colon tumor	universal normal control
	PRO21340	prostate tumor	universal normal control
	PRO21384	colon tumor	universal normal control
	PRO21384	lung tumor	universal normal control
	PRO21384	breast tumor	universal normal control

EXAMPLE 21: Tissue Expression Distribution

Oligonucleotide probes were constructed from the PRO polypeptide-encoding nucleotide sequence shown in the accompanying figures for use in quantitative PCR amplification reactions. The oligonucleotide probes were chosen so as to give an approximately 200-600 base pair amplified fragment from the 3' end of its associated template in a standard PCR reaction. The oligonucleotide probes were employed in standard quantitative PCR amplification reactions with cDNA libraries isolated from different human adult and/or fetal tissue sources and analyzed by agarose gel electrophoresis so as to obtain a quantitative determination of the level of expression of the PRO polypeptide-encoding nucleic acid in the various tissues tested. Knowledge of the expression pattern or the differential expression of the PRO polypeptide-encoding nucleic acid in various different human tissue types provides a diagnostic marker useful for tissue typing, with or without other tissue-specific markers, for determining the primary tissue source of a metastatic tumor, and the like. The results of these assays demonstrated the following:

(1) the DNA94849-2960 molecule is significantly expressed in the following tissues: cartilage, testis, colon tumor, heart, placenta, bone marrow, adrenal gland, prostate, spleen aortic endothelial cells and uterus, and not significantly expressed in the following tissues: HUVEC.

(2) the DNA100272-2969 molecule is significantly expressed in cartilage, testis, human umbilical vein endothelial cells (HUVEC), colon tumor, heart, placenta, bone marrow, adrenal gland, prostate, spleen and aortic endothelial cells; and not significantly expressed in uterus. Among a panel of normal and tumor cells examined, the DNA100272-2969 was found to be expressed in normal esophagus, esophageal tumor, normal stomach, stomach tumor, normal kidney, kidney tumor, normal lung, lung tumor, normal rectum, rectal tumor, normal liver and liver tumor.

(3) the DNA108696-2966 molecule is highly expressed in prostate and also expressed in testis, bone marrow and spleen. The DNA108696-2966 molecule is expressed in normal stomach, but not expressed in stomach tumor. The DNA108696-2966 molecule is not expressed in normal kidney, kidney tumor, normal lung, or lung tumor. The DNA108696-2966 molecule is highly expressed in normal rectum, lower expression in rectal tumor. The DNA108696-2966 molecule is not expressed in normal liver or liver tumor. The DNA108696-2966 molecule is not expressed in normal esophagus, esophageal tumor, cartilage, HUVEC, colon tumor, heart, placenta, adrenal gland, aortic endothelial cells and uterus.

(4) the DNA119498-2965 molecule is significantly expressed in the following tissues: highly expressed in aortic endothelial cells, and also significantly expressed in cartilage, testis, HUVEC, colon tumor, heart, placenta, bone marrow, adrenal gland, prostate and spleen. It is not significantly expressed in uterus.

(5) the DNA119530-2968 molecule is expressed in the following tissues: normal esophagus and not expressed in the following tissues: esophageal tumors, stomach tumors, normal stomach, normal kidney, kidney tumor, normal lung, lung tumor, normal rectum, rectal tumors, normal liver or liver tumors.

(6) the DNA129794-2967 molecule is significantly expressed in testis and adrenal gland; and not significantly expressed in cartilage, human umbilical vein endothelial cells (HUVEC), colon tumor, heart, placenta, bone marrow, prostate, spleen, aortic endothelial cells and uterus.

(7) the DNA131590-2962 molecule is significantly expressed in the following tissues: bone marrow, adrenal

gland, prostate, spleen, uterus, cartilage, testis, colon tumor, heart, and placenta, and not significantly expressed in the following tissues: HUVEC, and aortic endothelial cells.

(8) the DNA149995-2871 molecule is highly expressed in testis, and adrenal gland; expressed in cartilage, human umbilical vein endothelial cells (HUVEC), colon tumor, heart, prostate and uterus; weakly expressed in bone marrow, spleen and aortic endothelial cells; and not significantly expressed in placenta.

5 (9) the DNA167678-2963 molecule is significantly expressed in the following tissues: normal esophagus, esophageal tumor, highly expressed in normal stomach, stomach tumor, highly expressed in normal kidney, kidney tumor, expressed in lung, lung tumor, normal rectum, rectal tumor, weakly expressed in normal liver, and not significantly expressed in liver tumor.

10 (10) the DNA168028-2956 molecule is highly expressed in bone marrow; expressed in testis, human umbilical vein endothelial cells (HUVEC), colon tumor, heart, placenta, adrenal gland, prostate, spleen, aortic endothelial cells and uterus; and is weakly expressed in cartilage. Among a panel of normal and tumor samples examined, the DNA168028-2956 was found to be expressed in stomach tumor, normal kidney, kidney tumor, lung tumor, normal rectum and rectal tumor; and not expressed in normal esophagus, esophageal tumor, normal stomach, normal lung, normal liver and liver tumor.

15 (11) the DNA176775-2957 molecule is highly expressed in testis; expressed in cartilage and prostate; weakly expressed in adrenal gland, spleen and uterus; and not significantly expressed in human umbilical vein endothelial cells (HUVEC), colon tumor, heart, placenta, bone marrow and aortic endothelial cells.

20 (12) the DNA177313-2982 molecule is significantly expressed in prostate and aortic endothelial cells; and not significantly expressed in cartilage, testis, human umbilical vein endothelial cells (HUVEC), colon tumor, heart, placenta, bone marrow, adrenal gland, spleen and uterus. Among a panel of normal and tumor cells, the DNA177313-2982 molecule was found to be expressed in esophageal tumor but not in normal esophagus, normal stomach, stomach tumor, normal kidney, kidney tumor, normal lung, lung tumor, normal rectum, rectal tumor, normal liver and liver tumor.

WHAT IS CLAIMED IS:

1. Isolated nucleic acid having at least 80% nucleic acid sequence identity to a nucleotide sequence that encodes an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16),
5 Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54),
10 Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92),
15 Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 20 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), and Figure 244 (SEQ ID NO:244).

2. Isolated nucleic acid having at least 80% nucleic acid sequence identity to a nucleotide sequence selected from the group consisting of the nucleotide sequence shown in Figures 1A-1B (SEQ ID NO:1), Figure 3 (SEQ ID NO:3), Figure 5 (SEQ ID NO:5), Figure 7 (SEQ ID NO:7), Figure 9 (SEQ ID NO:9), Figure 11 (SEQ ID NO:11), Figure 13 (SEQ ID NO:13), Figure 15 (SEQ ID NO:15), Figure 17 (SEQ ID NO:17), Figure 19 (SEQ ID NO:19), Figure 21 (SEQ ID NO:21), Figure 23 (SEQ ID NO:23), Figure 25 (SEQ ID NO:25),
5 Figure 27 (SEQ ID NO:27), Figure 29 (SEQ ID NO:29), Figure 31 (SEQ ID NO:31), Figure 33 (SEQ ID NO:33), Figure 35 (SEQ ID NO:35), Figure 37 (SEQ ID NO:37), Figure 39 (SEQ ID NO:39), Figure 41 (SEQ ID NO:41), Figure 43 (SEQ ID NO:43), Figure 45 (SEQ ID NO:45), Figure 47 (SEQ ID NO:47), Figure 49 (SEQ ID NO:49), Figure 51 (SEQ ID NO:51), Figure 53 (SEQ ID NO:53), Figure 55 (SEQ ID NO:55), Figure 57 (SEQ ID NO:57), Figures 59A-59B (SEQ ID NO:59), Figure 61 (SEQ ID NO:61), Figure 63 (SEQ ID NO:63), Figure 65 (SEQ ID NO:65), Figure 67 (SEQ ID NO:67), Figure 69 (SEQ ID NO:69), Figure 71 (SEQ ID NO:71), Figure 73 (SEQ ID NO:73), Figure 75 (SEQ ID NO:75), Figure 77 (SEQ ID NO:77), Figure 79 (SEQ ID NO:79), Figure 81 (SEQ ID NO:81), Figure 83 (SEQ ID NO:83), Figure 85 (SEQ ID NO:85), Figure 87 (SEQ ID NO:87), Figure 89 (SEQ ID NO:89), Figure 91 (SEQ ID NO:91), Figure 93 (SEQ ID NO:93), Figure 95 (SEQ ID NO:95), Figure 97 (SEQ ID NO:97), Figure 99 (SEQ ID NO:99), Figure 101 (SEQ ID NO:101), Figure 103 (SEQ ID NO:103), Figure 105 (SEQ ID NO:105), Figure 107 (SEQ ID NO:107), Figure 109 (SEQ ID NO:109), Figure 111 (SEQ ID NO:111), Figure 113 (SEQ ID NO:113), Figure 115 (SEQ ID NO:115), Figure 117 (SEQ ID NO:117), Figure 119 (SEQ ID NO:119), Figure 121 (SEQ ID NO:121), Figure 123 (SEQ ID NO:123), Figure 125 (SEQ ID NO:125), Figure 127 (SEQ ID NO:127), Figure 129 (SEQ ID NO:129), Figure 131 (SEQ ID NO:131), Figure 133 (SEQ ID NO:133), Figure 135 (SEQ ID NO:135), Figure 137 (SEQ ID NO:137), Figure 139 (SEQ ID NO:139), Figure 141 (SEQ ID NO:141), Figure 143 (SEQ ID NO:143), Figure 145 (SEQ ID NO:145), Figure 147 (SEQ ID NO:147), Figure 149 (SEQ ID NO:149), Figure 151 (SEQ ID NO:151), Figure 153 (SEQ ID NO:153), Figure 155 (SEQ ID NO:155), Figure 157 (SEQ ID NO:157), Figure 159 (SEQ ID NO:159), Figure 161 (SEQ ID NO:161), Figure 163 (SEQ ID NO:163), Figure 165 (SEQ ID NO:165), Figure 167 (SEQ ID NO:167), Figure 169 (SEQ ID NO:169), Figure 171 (SEQ ID NO:171), Figure 173 (SEQ ID NO:173), Figure 175 (SEQ ID NO:175), Figure 177 (SEQ ID NO:177), Figure 179 (SEQ ID NO:179), Figure 181 (SEQ ID NO:181), Figure 183 (SEQ ID NO:183), Figure 185 (SEQ ID NO:185), Figure 187 (SEQ ID NO:187), Figure 189 (SEQ ID NO:189), Figure 191 (SEQ ID NO:191), Figure 193 (SEQ ID NO:193), Figure 195 (SEQ ID NO:195), Figure 197 (SEQ ID NO:197), Figure 199 (SEQ ID NO:199), Figure 201 (SEQ ID NO:201), Figure 203 (SEQ ID NO:203), Figure 205 (SEQ ID NO:205), Figure 207 (SEQ ID NO:207), Figure 209 (SEQ ID NO:209), Figure 211 (SEQ ID NO:211), Figure 213 (SEQ ID NO:213), Figure 215 (SEQ ID NO:215), Figure 217 (SEQ ID NO:217), Figure 219 (SEQ ID NO:219), Figure 221 (SEQ ID NO:221), Figure 223 (SEQ ID NO:223), Figure 225 (SEQ ID NO:225), Figure 227 (SEQ ID NO:227), Figure 229 (SEQ ID NO:229), Figure 231 (SEQ ID NO:231), Figure 233 (SEQ ID NO:233), Figure 235 (SEQ ID NO:235), Figure 237 (SEQ ID NO:237), Figure 239 (SEQ ID NO:239), Figure 241 (SEQ ID NO:241), and Figure 243 (SEQ ID NO:243).

3. Isolated nucleic acid having at least 80% nucleic acid sequence identity to a nucleotide sequence

selected from the group consisting of the full-length coding sequence of the nucleotide sequence shown in Figures 1A-1B (SEQ ID NO:1), Figure 3 (SEQ ID NO:3), Figure 5 (SEQ ID NO:5), Figure 7 (SEQ ID NO:7), Figure 9 (SEQ ID NO:9), Figure 11 (SEQ ID NO:11), Figure 13 (SEQ ID NO:13), Figure 15 (SEQ ID NO:15), Figure 17 (SEQ ID NO:17), Figure 19 (SEQ ID NO:19), Figure 21 (SEQ ID NO:21), Figure 23 (SEQ ID NO:23), Figure 25 (SEQ ID NO:25), Figure 27 (SEQ ID NO:27), Figure 29 (SEQ ID NO:29), Figure 31 (SEQ ID NO:31), Figure 33 (SEQ ID NO:33), Figure 35 (SEQ ID NO:35), Figure 37 (SEQ ID NO:37), Figure 39 (SEQ ID NO:39), Figure 41 (SEQ ID NO:41), Figure 43 (SEQ ID NO:43), Figure 45 (SEQ ID NO:45), Figure 47 (SEQ ID NO:47), Figure 49 (SEQ ID NO:49), Figure 51 (SEQ ID NO:51), Figure 53 (SEQ ID NO:53), Figure 55 (SEQ ID NO:55), Figure 57 (SEQ ID NO:57), Figures 59A-59B (SEQ ID NO:59), Figure 61 (SEQ ID NO:61), Figure 63 (SEQ ID NO:63), Figure 65 (SEQ ID NO:65), Figure 67 (SEQ ID NO:67), Figure 69 (SEQ ID NO:69), Figure 71 (SEQ ID NO:71), Figure 73 (SEQ ID NO:73), Figure 75 (SEQ ID NO:75), Figure 77 (SEQ ID NO:77), Figure 79 (SEQ ID NO:79), Figure 81 (SEQ ID NO:81), Figure 83 (SEQ ID NO:83), Figure 85 (SEQ ID NO:85), Figure 87 (SEQ ID NO:87), Figure 89 (SEQ ID NO:89), Figure 91 (SEQ ID NO:91), Figure 93 (SEQ ID NO:93), Figure 95 (SEQ ID NO:95), Figure 97 (SEQ ID NO:97), Figure 99 (SEQ ID NO:99), Figure 101 (SEQ ID NO:101), Figure 103 (SEQ ID NO:103), Figure 105 (SEQ ID NO:105), Figure 107 (SEQ ID NO:107), Figure 109 (SEQ ID NO:109), Figure 111 (SEQ ID NO:111), Figure 113 (SEQ ID NO:113), Figure 115 (SEQ ID NO:115), Figure 117 (SEQ ID NO:117), Figure 119 (SEQ ID NO:119), Figure 121 (SEQ ID NO:121), Figure 123 (SEQ ID NO:123), Figure 125 (SEQ ID NO:125), Figure 127 (SEQ ID NO:127), Figure 129 (SEQ ID NO:129), Figure 131 (SEQ ID NO:131), Figure 133 (SEQ ID NO:133), Figure 135 (SEQ ID NO:135), Figure 137 (SEQ ID NO:137), Figure 139 (SEQ ID NO:139), Figure 141 (SEQ ID NO:141), Figure 143 (SEQ ID NO:143), Figure 145 (SEQ ID NO:145), Figure 147 (SEQ ID NO:147), Figure 149 (SEQ ID NO:149), Figure 151 (SEQ ID NO:151), Figure 153 (SEQ ID NO:153), Figure 155 (SEQ ID NO:155), Figure 157 (SEQ ID NO:157), Figure 159 (SEQ ID NO:159), Figure 161 (SEQ ID NO:161), Figure 163 (SEQ ID NO:163), Figure 165 (SEQ ID NO:165), Figure 167 (SEQ ID NO:167), Figure 169 (SEQ ID NO:169), Figure 171 (SEQ ID NO:171), Figure 173 (SEQ ID NO:173), Figure 175 (SEQ ID NO:175), Figure 177 (SEQ ID NO:177), Figure 179 (SEQ ID NO:179), Figure 181 (SEQ ID NO:181), Figure 183 (SEQ ID NO:183), Figure 185 (SEQ ID NO:185), Figure 187 (SEQ ID NO:187), Figure 189 (SEQ ID NO:189), Figure 191 (SEQ ID NO:191), Figure 193 (SEQ ID NO:193), Figure 195 (SEQ ID NO:195), Figure 197 (SEQ ID NO:197), Figure 199 (SEQ ID NO:199), Figure 201 (SEQ ID NO:201), Figure 203 (SEQ ID NO:203), Figure 205 (SEQ ID NO:205), Figure 207 (SEQ ID NO:207), Figure 209 (SEQ ID NO:209), Figure 211 (SEQ ID NO:211), Figure 213 (SEQ ID NO:213), Figure 215 (SEQ ID NO:215), Figure 217 (SEQ ID NO:217), Figure 219 (SEQ ID NO:219), Figure 221 (SEQ ID NO:221), Figure 223 (SEQ ID NO:223), Figure 225 (SEQ ID NO:225), Figure 227 (SEQ ID NO:227), Figure 229 (SEQ ID NO:229), Figure 231 (SEQ ID NO:231), Figure 233 (SEQ ID NO:233), Figure 235 (SEQ ID NO:235), Figure 237 (SEQ ID NO:237), Figure 239 (SEQ ID NO:239), Figure 241 (SEQ ID NO:241), and Figure 243 (SEQ ID NO:243).

35

4. Isolated nucleic acid having at least 80% nucleic acid sequence identity to the full-length coding sequence of the DNA deposited under any ATCC accession number shown in Table 7.

5. A vector comprising the nucleic acid of Claim 1.
6. A host cell comprising the vector of Claim 5.
7. The host cell of Claim 6, wherein said cell is a CHO cell.
- 5 8. The host cell of Claim 6, wherein said cell is an *E. coli*.
9. The host cell of Claim 6, wherein said cell is a yeast cell.
- 10 10. A process for producing a PRO polypeptide comprising culturing the host cell of Claim 6 under conditions suitable for expression of said PRO polypeptide and recovering said PRO polypeptide from the cell culture.
11. An isolated polypeptide having at least 80% amino acid sequence identity to an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure

166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), and Figure 244 (SEQ ID NO:244).

12. An isolated polypeptide having at least 80% amino acid sequence identity to an amino acid sequence encoded by the full-length coding sequence of the DNA deposited under any ATCC accession number 15 shown in Table 7.

13. A chimeric molecule comprising a polypeptide according to Claim 11 fused to a heterologous amino acid sequence.

20 14. The chimeric molecule of Claim 13, wherein said heterologous amino acid sequence is an epitope tag sequence.

25 15. The chimeric molecule of Claim 13, wherein said heterologous amino acid sequence is a Fc region of an immunoglobulin.

16. An antibody which specifically binds to a polypeptide according to Claim 11.

30 17. The antibody of Claim 16, wherein said antibody is a monoclonal antibody, a humanized antibody or a single-chain antibody.

18. Isolated nucleic acid having at least 80% nucleic acid sequence identity to:

(a) a nucleotide sequence encoding the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ

ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ 5 ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 10 ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID 15 NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 20 ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID 25 NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), or Figure 244 (SEQ ID NO:244), lacking its associated signal peptide;

(b) a nucleotide sequence encoding an extracellular domain of the polypeptide shown in Figure 2 30 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID 35 NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure

64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70),
Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID
NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ
ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94
(SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100),
5 Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ
ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure
116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID
NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure
130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID
10 NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure
144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID
NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure
158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID
NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure
15 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID
NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure
186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID
NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure
200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID
20 NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure
214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID
NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure
228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID
NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure
25 242 (SEQ ID NO:242), or Figure 244 (SEQ ID NO:244), with its associated signal peptide; or
(c) a nucleotide sequence encoding an extracellular domain of the polypeptide shown in Figure 2
(SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ
ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18
(SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure
30 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32),
Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID
NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ
ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56
(SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure
35 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70),
Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID
NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ

ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), or Figure 244 (SEQ ID NO:244), lacking its associated signal peptide.

19. An isolated polypeptide having at least 80% amino acid sequence identity to:

25 (a) an amino acid sequence of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ

ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), or Figure 244 (SEQ ID NO:244), lacking its associated signal peptide;

(b) an amino acid sequence of an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ

ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), or Figure 244 (SEQ ID NO:244), with its associated signal peptide; or

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure

138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), or Figure 244 (SEQ ID NO:244), lacking its associated signal peptide.

20. A method for stimulating the proliferation of or gene expression in pericyte cells, said method comprising contacting said cells with a PRO982, PRO1160, PRO1187, or PRO1329 polypeptide, wherein the proliferation of or gene expression in said cells is stimulated.

20

21. A method for stimulating the proliferation or differentiation of chondrocyte cells, said method comprising contacting said cells with a PRO357, PRO229, PRO1272 or PRO4405 polypeptide, wherein the proliferation or differentiation of said cells is stimulated.

25

22. A method for stimulating the release of TNF- α from human blood, said method comprising contacting said blood with a PRO231, PRO357, PRO725, PRO1155, PRO1306 or PRO1419 polypeptide, wherein the release of TNF- α from said blood is stimulated.

30

23. A method for stimulating the proliferation of normal human dermal fibroblast cells, said method comprising contacting said cells with a PRO982, PRO357, PRO725, PRO1306, PRO1419, PRO229, PRO1272, PRO181, PRO214, PRO247, PRO337, PRO526, PRO363, PRO531, PRO1083, PRO840, PRO1080, PRO788, PRO1478, PRO1134, PRO826, PRO1005, PRO809, PRO1194, PRO1071, PRO1411, PRO1309, PRO1025, PRO1181, PRO1126, PRO1186, PRO1192, PRO1244, PRO1274, PRO1412, PRO1286, PRO1330, PRO1347, PRO1305, PRO1273, PRO1279, PRO1340, PRO1338, PRO1343, PRO1376, PRO1387, PRO1409, PRO1488, PRO1474, PRO1917, PRO1760, PRO1567, PRO1887, PRO1928, PRO4341, PRO5723, PRO1801, PRO4333, PRO3543, PRO3444, PRO4302, PRO4322, PRO5725, PRO4408, PRO9940, PRO7154, PRO7425, PRO6079, PRO9836 or PRO10096 polypeptide, wherein the proliferation of said cells is stimulated.

24. A method for detecting the presence of tumor in an mammal, said method comprising comparing the level of expression of any PRO polypeptide shown in Table 8 in (a) a test sample of cells taken from said mammal and (b) a control sample of normal cells of the same cell type, wherein a higher level of expression of said PRO polypeptide in the test sample as compared to the control sample is indicative of the presence of tumor in said mammal.

5

25. The method of Claim 24, wherein said tumor is lung tumor, colon tumor, breast tumor, prostate tumor, rectal tumor, or liver tumor.

26. An oligonucleotide probe derived from any of the nucleotide sequences shown in the
10 accompanying figures.

FIGURE 1A

GCAGCCCTAGCAGGG**ATGG**GACATGATGCTGTTGGCAGGGTCTTGTGCTCGAACAGTG
GCTGGCGCGGTGCTCCTCAGCCTGTGCTGCCTGCTACCCCTGCCTCCCGCTGGACAGA
GTGTGGACTTCCCTGGCGCCGTGGACAACATGATGGTCAGAAAAGGGGACACGGCGGTG
CTTAGGTGTTATTGGAAGATGGAGCTCAAAGGGTGCCTGGCTGAACCGGTCAAGTATTAT
TTTGCAGGGAGGTGATAAGTGGTCAGTGGATCCTCGAGTTCAATTCAACATTGAATAAAA
GGGACTACAGCCTCCAGATACAGAATGTAGATGTGACAGATGATGGCCCACACGTGTTCT
GTCAGACTCAACATACACCCAGAACAAATGCAAGGTGATCTAAGTGTGCAAGTCCCTCAA
GATATATGACATCTCAAATGATATGACCGTCAATGAAGGAACCAACGTCACTTTACTTGT
TGGCCACTGGAAACCAAGAGCCTTCCATTCTTGGCGACACATCTCCCCATCAGCAAAACCA
TTTGAAGGATGGACAATATTGGACATTGAAATTACAAGGGACCAGGCTGGGAATATGA
ATGCAGTGCAGGAAATGATGTGTCATTCCAGATGTGAGGAAAGTAAAGTTGTTGTCAACT
TTGCTCCTACTATTCAAGGAAATTAAATCTGGCACCGTGACCCCCGGACGCAGTGGCCTGATA
AGATGTGAAGGGCAGGTGTGCCCTCAGCCTTGAATGGTACAAGGGAGAGAAGAAGCT
CTTCATGGCCAACAAGGAATTATTCAAATTTAGCACAAGATCCATTCACTGTTA
CCAACGTGACACAGGAGCACTCGGCAATTATACTTGTGTCGGCTGCCAACAAAGCTAGGCACA
ACCAATGCGAGCCTGCCTCTAACCTCCAAGTACAGCCCAGTATGGAATTACCGGGAGCAG
TGATGTTCTTCTCCTGCTGGTACCTGTGTTGACACTGTCCTCTTACCCAGCATATTCT
ACCTGAAGAATGCCATTCAAA**TAA**ATTCAAAGACCCATAAAAGGCTTTAAGGATTCTCT
GAAAGTGCTGATGGCTGGATCCAATCTGGTACAGTTGTTAAAGCAGCGTGGATATAATC
AGCAGTGCTTACATGGGATGATCGCCTCTGTAGAATTGCTCATTATGTAATACCTTAAAT
TCTACTCTTTGATTAGCTACATTACCTTGTGAAAGCAGTACACATTGTCCTTTTTAAG
ACGTGAAAGCTCTGAAATTACTTTAGGGATTTAATTGTGATTTCATGTTGTAATCTAC
AACTTTCAAAAGCATTCAAGTACGGTCTGCTAGGTTGCAAGGCTGTAGTTACAAAAACGAA
TATTGCAAGTGAATATGTGATTCTTAAGGCTGCAATACAAGCATTCAAGTCCCTGTTCAAT
AAGAGTCATCCACATTACAAGATGCAATTCTTTCTTTGATAAAAAAGCAAATAATA
TTGCCTTCAGATTATTCTCAAATATAACACATATCTAGATTCTGCTCGCATGATAT
TCAGGTTTCAGGAATGAGCCTTGTAAATAACTGGCTGTGCAAGCTCTGCTCTTCTGT
AAGTTCAAGCATGGGTGTGCCTTCATACATAATTCTCTTGTCTCCAACATAATAAA
AATGTTTGCTAAATCTACAATTGAAAGTAAAATAACCAAGAGTGTCAAGTTAACCA
TACACTATCTCTAAAGTAACGAAGGAGCTATTGGACTGTAAAAATCTCTGCACAGCAA
TGGGGTTGAGAATTGCCCCACACTAACTCAGTTGTGATGAGAGACAATTAAAC
AGTATAGTAAATATAACCATATGATTCTTAGTTGAGCTAAATGTTAGATCCACCGTGGGA
AATCATTCCCTTAAATGACAGCACAGTCCACTCAAAGGATTGCCTAGCAATACAGCATCT
TTCCCTTCACTAGTCCAAGCCAAAATTAAAGATGATTGTGAGAAAGGGCACAAAGTCC
TATCACCTAATTACAAGAGTTGGTAAGCGCTCATCATTAAATTGTTATTGTGGCAGCTAA
GTTAGTATGACAGAGGCAGTGCCTGTGGACAGGGCATTGCAATTCTCATCTGAAA
GTATCACTCAGTTGATAGTCTGGAATGCACTGTTATATATTAAAATCTCAAATATATTA
TAACAAACATTCTATATCGGTATGTAGCAGACCAATCTCTAAATAGCTAATTCTCAATAAA
AATCTTCTATATAGCCATTCACTGCAACAAAGTAAAATCAAAAGACCATCTTATT
TTCCTTACATGATATATGTAAGATGCGATCAAATAAGACAAAACCCAGTGTGAGAATAT
CTTAAGATAAGTAATTATCAAATTATTGTGAATGTTAAATTATTCCTACTATAAGAAGCAA
AACTACATTGGAGGAAATGCTGTTACTCTAACATTAAATTACAGGAATAGTTGATGG
TTTCACTCTTACTAAAGAAAGGCCATCACCTGAAAGCCATTTCAGGTTGATGAAGTT
ACCAATTTCAGTACACCTAAATTCTACAAATAGTCCCCTTACAAGTTGTAACAACAAAG
ACCCCTATAAAATAGATACAAGAAATTGCGAGTGGTTACATATTGAGATATCTAG
TATGTTGCCCTAGCAGGGATGGCTAAAAACTGTGATTGTTCTTCAAGTAAAACCTAGT
CCCAAAGTACATCATAAAATCAATTAAATTAGAAAAATGAATCTTAAATGAGGGGACATAAG
TATACTCTTCCACAAAATGGCAATAATAAGGCATAAGCTAGTAAATCTACTAACTGTAAT
AAATGTATGACATTATTGATTGATACATTAAAGAGTTTGAACAAATATGGCATT
TAACTTTATTATTGCTTTAAGAAATTCTTGTGGAATTGTTGAATAAAACTATAAA
AATATTATTGTATTGCAGCTTAAAGTGGCACACTCCATAATAATCTACTTAGAAAT

FIGURE 1B

AGTGGTGCTACCACAAAAATGTTAACCATCAGTACCAATTGTTGGGAGAAAGAAACAGATC
AAGAATGCATATTATTCACTGACCGCTTCCTAGAGTTAAACACCTCCTCTTGTAAAGGTT
TGTAGGTAAATTGAGGTATAAACTATGGATGAACCAAATAATTAGTTCAAAGTGTGTCATG
ATTCCAAATTGTTGGAGTCTGGTCTTACCATAGAATGTGACAGAAGTACAGTCAGCT
CAGTAGCTATATGTATTGCCTTATGTAGAAGAGACTTCTTGAGTGACATTAAATA
GAGGAGGTATTCACTATGTTCTGTATCACAGCAGCATTCTAGTCCTAGGCCCTCGGA
CAGAGTGAATCATGAGTATTATGAGTCATATTGCAAAATAAGGCTACAGTATTGCTT
TTTGTGTGAATGTATTGCATATAATGTCAAGTAGATGATTACATTATGGACATATAA
AATGTCTGATTACCCATTATCAGTCCTGACTGTACAAGATTGTCAATTTCAGAATAG
CAGTTTATAAATTGATTATCTTTAATCTATAACAATTGTGTTAGCTGTTATTTCAGG
ANTATATTCTACAAGTCCACTTGTGGACTCCTTGTGCCCCTATTTTTAAAG
AAGGAAGAAAGAAAAATAAGTAGCAGTTAAAATGAGAATGGAGAGAAAAGAAAAGAATG
AAAAGGAAAGGCAGTAAAGAGGGAAAAAAAGGAAGGGATGGAAGGAATGAAGGAAGGAAGGG
AGGAAGGGAGAAGGTAGGAAGAAAGAAAGGATGAGAGGGAGGAAGAATCAGAGTATTAGG
GTAGTTAACTTACACATTGCAATTCTTAGTTAACTGCAAGTGGTGTAACTATGTTTCAA
TGATCGCATTTGAAACATAAGTCCTATTATACCATTAAGTCCTATTATGCAGCAATTATAT
AATAAAAGTACTGCCAAGTTAGTAATGTGGGTGTTTGAGACACTAAAGATTTGAG
AGGGAGAATTCAAACCTAAAGCCACTTTGGGGGTTATAACTTAAGTAACTGAAAAATTAAAC
CTTCATCATAACATTAAAGCTATCTAGAAAGTAGACTGGAGAACTGAGAAAATTACCCAG
GTAATTCAAGGGAAAAAAATATATATATAAATACCCCTACATTGAAGTCAGAAA
ACTCTGAAAAACTGAATTATCAAAGTCATCTATAATGATCAAATTACTGAACAAATTG
TTAATTATCCATTGTGCTTAGCTTGTGACACAGCCAAAGTTACCTATTAACTTTCA
ATAAAAATTGTTTGTGAAATCCAGAAATGATTAAAAAGAGGTAGGTTAACTATTAA
TTGAAGTATGTGGATGTACAGTATTCAATAGATATGAATAATGGTATGCCTTAA
GATTCTTGAATATGTATTCAATAGACTGGAAAAAGCTTCTGTCTTTAGTAAA
CATCCATATTCTACACCTGATGAAATATGTTGACTGTTCCAATAGGTGAATATAAAC
TCAGTTATCAATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 2

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA92259
><subunit 1 of 1, 354 aa, 1 stop
><MW: 38719, pI: 6.12, NX(S/T): 6
MDMMLLVQGACCSNQWLAALSLCCLLPSCLPAGQSVDFPWAADVNMVRKGDTAVLRCYL
EDGASKGAWLNRSSIIFAGGDKWSVDPRVSISTLNKRDYSLQIQNVDVTDDGPYTCVQTQH
TPRTMQVHLTVQVPPKIYDISNDMTVNEGTVTLTCLATGKPEPSISWRHISPSAKPFENGQ
YLDIYGITRDQAGEYECASAENDVSFPDVRKVVVNFAPTIQEIKSGTVPGRSGLIRCEGA
GVPPPAFEWYKGEKKLFNGQQGIIIQNFSTRSILTVTNVTQEHFGNYTCVAANKLGTNASL
PLNPPSTAQYGITGSADVLFSCWYLVLTSSFTSIFYLKNAILQ

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 322-343

N-glycosylation sites.

amino acids 73-77, 155-159, 275-279, 286-290, 294-298, 307-311

Tyrosine kinase phosphorylation site.

amino acids 180-188

N-myristoylation sites.

amino acids 9-15, 65-71, 69-75, 153-159, 241-247, 293-299,
304-310, 321-327

Myelin P0 protein.

amino acids 94-123

FIGURE 3

CACTGCCCGTCCGCTTCAGCAGCCGGCGGGCGGTGGAAAAGCGAGTGAAGAGAGAGCGC
GACGGCGGCGGCGGCGGCGAGCTATTGCTGGACGCCAGTGGAGAGAGCGAGGCCTGAG
CCTCTGCGTCTAGGATCAAATGGTTCAATCCAGAATACTATGAAGGCAAGAACGTCCTC
CTCACAGGAGCTACCGGTTCTAGGGAGGTGCTCTGGAAAAGTGTGAGGCTTGTCC
TAAGGTGAATTCACTATGTTGGTGGAGGCAGAAAGCTGGACAGACACCACAAGAGCGAG
TGGAAAGAAGTCCTAGTGGCAAGCTTTGACAGATTGAGAGATGAAAATCCAGATTTAGA
GAGAAAATTATAGCAATCACACAGCGAACTCACCCAAACCTAAACTGGCTCTCAGTGAAGAAGA
TAAAGAGGTGATCATAGATTCTACCAATATTATTCACGTGAGCTACAGTAAGGTTA
ATGAAAATTAAAGAGATGCTGTTAGTTAAATGTGATTGCAACGCGACAGCTTATTCTCCT
GCACAAACAAATGAAGAATCTGGAAAGTGTTCATGCATGTATCAACAGCATATGCCTACTGTAA
TCGCAAGCATATTGATGAAGTAGTCTATCCACCACTGTGGATCCCAAAGAAGCTGATTGATTCT
TTAGAGTGGATGGATGATGCCCTAGTAAATGATATCACGCCAAATTGATAGGAGACAGACC
TAATACATACATATAACACAAAAGCATTGGCAGAATATGTTGTACAACAAGAAGGAGCAAAC
TAAATGTGGCAATTGTAAGGCCATCGATTGTTGGTGCAGTTGGAAAGAACCTTTCAGGA
TGGATTGATAACCTTAATGGACCAAGTGGTCTCTTATTGCGGCAGGGAAAGGAATTCTCG
AACAAATACGTGCCCTCAACAATGCCCTGCAGATCTGTTCTGTAGATGTAGTTGTCACAA
TGAGTCCTGCCGAGCCTGGTATTCCGGAGTTAATAGACCAAGAACATCATGGGTGATAAT
TGTACAACAGGCAGCACTAACCTTCCACTGGGTGAAGTTGAGTACCATGTAATTCCAC
TTTCAAGAGGAATCCTCTCGAACAGGCCCTCAGACGCCAATGTAATCTAACCTCCAATC
ATCTTTATATCATTACTGGATTGCTGTAAGCCATAAGGCCAGCATTCTGTATGATATC
TACCTCAGGATGACTGGAAGAACGCCAAGGATGATGAAAACAATAACTCGTCTTCACAAAGC
TATGGTGTCTGAATATTCAAGTAATTCTGGGTTGGAACTGAGAATGTCATA
TGTAAATGAATCAACTAAACCTGAAGATAAAAGACCTCAATATTGATGTACGGCAGTTA
CATTGGCAGAATATAGAGAACTACTGCTGGAACTAAGAAGTACGTATTGAATGAAGA
AATGTCGGCTCCCTGCAGCCAGAAAACATCTGAACAAAGTTGCGGAATATACGTTATGGTT
TTAATACTATCCTGTGATCCTCATCTGGCGATTTTATTGCAAGATCACAAATGGCAAGA
AATATCTGGTACCTTGTGGTTAGTCGTGTTACAAGTTTGTCAACTTCCGAGCATCCAG
CACTATGAGATACTGAAGACCAAGGATTGAGCATTAGAACATCTACATATGGTATCTAA
ATGTACAAAATGAAAATGTATAAGTCATCTCACTTTGTCAAGACATTAAACCATCTTAG
ATCGGAGTGTGAAGTAAATTATGGTATATTGTAAACATTAAATGTTATGCTCATAAA
ACTTAGTGAACACACTGTGTTATGCCAGCTAAATCTACAGTAGGCCACCAAAACCATGACTT
AATATTGTAGCCCTAGAAGAAAGGGGTGTGCTGAGGACAAGAGTGGGAAATAGGAACACT
GACCGATATAACTGTCAATTCTGGAACATATTAATTAAAATAATGCTTAACATATAGT
GAATTCTAATTCTAATGTCAGTGCATGGAAAGACATTATTTGGACAGTATAACTAGCAAA
GTTGGTAGATATTGATTCTCATTGTTTGTGTTTCAATTGTTGAAGTGGGTTTAGTT
TGTTAAAATTATAACCAGCGTATTTCACATCATTGTAAGTTAAATGATATCAAACATG
AAAGAGATGTTCTCATTTCTTCTGATTAACGCTGATGCATATCATTCTATAA
GTAATCAGTTGCTTTAAATCAGAAGGCTATTATCTAATGACCCATTGATCTAAAT
GGGTTGAGAATCCATATCAGCAACACATACGTGTTTGTGACAGAAAGTGAACAAATTCCG
TAAAATGTTAGTATCAAAAAGAATAGGAATACAGTTCTTCCACATTGATCAAATAAA
AATCTTGTGAGATTGTTAAAAA

5/246

FIGURE 4

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA94849
><subunit 1 of 1, 515 aa, 1 stop
><MW: 59357, pi: 9.40, NX(S/T): 3
MVSIPYYEGKNVLLTGATGFLGKVLEKLLRSCPKVNSVYVLVRQKAGQTQPERVEEVLSGKLF
DRLRDENPDFREKIIAINSELTQPKLALSEEDKEVIIDSTNIIFHCAATVRFNENLRDAVQLNVIATRQLILLQQMKNLEVF
MHVSTAYAYCNRKHIDEVYVPPVDPKKLIDSLEWMDDGLVNDITPKLIGDRPNTYIYT
KALAEYVQQEGAKLNVAIVRPSIVGASWKEPFPGWIDNFNGPSGLFIAAGKGILRTIRASNNALADLVPVDVVV
NMSLAAAWYSGVNRPRNIMVYNCTTGSTNPFHGEVEYHVISTFKRNPLEQAFRRPNVNLTSNHLLYHYWI
AVSHKAPAFLYDIYLRMTGRSPRMMKTITRLHKAMVLEYFTSN
SVWNTENVNMLMNQLNPEDKKTFNIDVRQLHWA
EYIENYCLGKKYVLNEEMSGLPAA
RKHLNKLRNIRYGFNTILVILIWRIFIARSQMARNI
WYFVVSLCYKFLSYFRASSTMRY
```

Important features of the protein:**Transmembrane domain:**

Amino acids 469-488

N-glycosylation sites:

Amino acids 283-287; 304-308; 341-345

Tyrosine kinase phosphorylation site:

Amino acids 160-169

N-myristoylation sites:

Amino acids 219-225; 252-258; 260-266; 452-458

Leucine zipper pattern:

Amino acids 439-461

FIGURE 5

7/246

FIGURE 6

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96883
><subunit 1 of 1, 514 aa, 1 stop
><MW: 55687, pI: 8.78, NX(S/T): 2
MPAVSGPGPLFCLLLLLDPHSPETGCPLRRFEYKLSFKGPRLLALPGAGIPFWSHHGDA
ILGLEEVRLTPSMRNRSGAVWSRASVPFSAWEVEVQMRVTGLGRRGAQGMAVWYTRGRGH
VGSVLGGLASWDGIGIFFDSPAEDTQDSPAIRVLASDGHIPSEQPGDGASQGLGSCHWDF
RNRPHSFRARITYWGQRLRMSLNSGLTPSDPGEFCVDVGPLLVPGGFFGVSAATGTLAG
EDPTGQVPPQPFLEMQQLRLARQLEGLWARLGLGTREDVTPKSDSEAQGEGERLFDLEET
LGRHRRILQALRGQLSKQLAQERQWKQQLGPPGQARPDGGAQDASCQIPSTPGRGGHLS
MSLNKDSAKVGALLHGQWTLLQALQEMRDAAVRMAAEAQVSYLPVGIEHHFELDHILGL
LQEELRGPAKAAAKAPRPPGQPPRASSCLQPGIFLFYLLIQTVGFFGYVHFRQELNKSLO
ECLSTGSLPLGPAPHTPRALGILRRQPLPASMPA
```

Important features of the protein:

Signal peptide:

Amino acids 1-23

Transmembrane domain:

Amino acids 215-232; 450-465

N-glycosylation sites:

Amino acids 75-79; 476-480

Glycosaminoglycan attachment site:

Amino acids 5-9

N-myristoylation sites:

Amino acids 78-84; 122-128; 126-132; 168-174; 172-178;
205-211; 226-232; 230-236; 236-242; 356-362

Amidation site:

Amino acids 102-106

FIGURE 7

9/246

FIGURE 8

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96894
><subunit 1 of 1, 361 aa, 1 stop
><MW: 40747, pI: 9.20, NX(S/T): 1
MAWQGWPAAWQWVAGCWLLVLVLLVSPRGCRARRGLRGMAHSQRLLFRIGYSLYT
RTWLGYLFYRQQLRRARNRYPKGHSKTQPRLFNGVKVLPIPVLSDNYSYLIIDTQAQLAV
AVDPSDPRAVQASIEKEGVTLVAILCTHKHWDHSGGNRDLSSRRHRDCRVYGSQDGIPYL
THPLCHQDVSVGRLQIRALATPGHTQGHLVYLLDGEPYKGPSCLFSGDLLFLSGCGRTF
EGNAETMLSLLDTVLGLGDDTLLWPGHEYAEENLGFAGVVEPENLARERKMQWVQRQRLE
RKGTCPSTLGEERSYNPFLRTHCLALQEALGPGPGPTGDDDSRAQLLEELRRLKDMHKS
K
```

Important features of the protein:**Signal peptide:**

Amino acids 1-35

N-glycosylation site:

Amino acids 106-110

Glycosaminoglycan attachment site:

Amino acids 234-238

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 301-305

Tyrosine kinase phosphorylation site:

Amino acids 162-171

N-myristoylation sites:

Amino acids 41-47;235-241;242-248;303-309

Prokaryotic membrane lipoprotein lipid attachment site:

Amino acids 6-17

cAMP phosphodiesterases class-II proteins:

Amino acids 144-161

10/246

FIGURE 9

GCTGACAATCCCCTTGACGTTCTATCCCGGAAGCTCCACCTGGGGCCCAATGTTGGCGTGA
TGTTCCCTCGCCTGTCTGCCTGGAAAACGGTCTCCAAAGCTCCACTGGCAGCCACTTCT
CCATGTTGGCATTGGAGACATCGTTATGCCTGGTCTCCTACTATGCTTGTCCCTCGCTAT
GACAACATACAAAAAGCAAGCCAGTGGGACTCCTGTGGGCCCCCTGGACCTGCCAACATCTC
CGGGCGCATGCAGAAGGTCTCCTACTCTCACTGCACCCCTCATCGGATACTTGTAGGCCTGC
TCACTGCTACTGTGGCGTCTCGCATTCACCGGGCCCGCCAGCCGCCCTCTCTATTGGTG
CCATTTACTTATTGCCACTCCTCACGATGGCCTATTAAAGGGCGACCTCCGGCGATGTG
GTCTGAGCCTTCCACTCCAAGTCCAGCAGCTCCGATTCCTGGAAGTATGATGGATCACGT
GGAAAGTGACCAGATGGCGTCATAGTCCTTTCTCAACTCATGGTTGTTCCCTTTAG
AGCTGGCCTGGTACTCAGAAATGTACCTGTGTTAAGGAACTGCCGTGTGACTGGATTGGC
ATTGAAAGGGAGCTCGTTGCAGGAGAGGGTGTGGAGCCTGTTGGTTCCCTCTCTCC
TGCAGATGTAGAGGTGGGGCCCCCTCAAGAGGGACAGGCCTCTCCCCAGCGCGCCCTCC
CCACGTTTATGGATCTGCACCACTGTTACCTCTGGGGAGATGGAGATTGACTGTT
TAAAAACTGAAAACAGCGAGGAGTCTTCTAGAACACTAAAGGATGAAAAAAT
TAGC

11/246

FIGURE 10

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA100272
><subunit 1 of 1, 108 aa, 1 stop
><MW: 12055, pI: 4.69, NX(S/T): 0
MMDHVESDQMAVIVLSQLMVCFLLELAWYSEMYLCLRNCRVTGFGIERELVCRREVLEP
CLVPSLPADVEVGPLPRGTGLSPARLPPTFLWICTRLLPSGGDGDLTV
```

Important features of the protein:

Signal peptide:

Amino acids 1-30

N-myristoylation site:

Amino acids 80-86

12/246

FIGURE 11

TCGCACACTGGTGGCTTCAGAAGAAATTCTCAACACCTAGCTGCCAGAGAGTCTATGTATG
GGATTGAACAATCTGTAAACTAAAGGATCTAATCATGAAAATAAGTATGATAAATTATAAG
TCACTATTGGCACTGTTGTTATATTAGCCTCTGGATCATTTACAGTTCCAGAACACTC
CACAAAGGTTGGTCTGCTCTAAACTATCCATCTCCCTCATTACTGGAACAACCTCACAA
AGTCCTTATTCCCTAAACACCACTGATATCATTAAAGCCACTAACAGAGACTGAACACTCAGA
ATAAAGGAAATCATAGAGAAACTAGATCAGCAGATCCCACCCAGACCTTCACCCACGTGAA
CACCACCACCGGCCACACATAGCACAGCCACCATCCTCAACCCTCGAGATACTGACTGCA
GGGGAGACCAGCTGCACATCCTGCTGGAGGTGAGGGACCACCTGGGACGCAGGAAGCAATAT
GGCGGGGATTCCTGAGGGCCAGGATGTTCCCAAGCGCTGATGGCAGGTGCTTCAGGAAA
GGTGA^TCTGACTCAACAACGGCACCTACCTGGTCAGCTTCACTCTGTTCTGGGAGGGCCAGG
TCTCTGTCTGCTGCTCATCCACCCAGTGAAGGGGTGTCAGCTCTGGAGTGCAAGG
AACCAAGGCTATGACAGGGT^TGATCTTCACTGGCCAGTTGTCATGGCACTTCCAAGTCCA
CTCTGAATGTGGCCTGATCCTAAACACAAATGCTGAATTGTGCCAGTACCTGGACAACAGAG
ACCAAGAAGGCTCTACTGTGTGAGGCC^TCAACACATGCCCTGTGCTGCACTCACTCACATG
TATTCTAAGAACAGAAAGTTCTTATCTTAGCAAACAAGAAAAGAGCCTTTGAAAGGTC
AAATGTGGGTGTAGAGATTATGAAAAATTCAATACAATTAGTGTCTCAAATGCAACAAAG
AAACAGTTGCAATGAAAGAGAAATGCAAGTTGGAATGACATCCACAAATCCCAGTGGC^TAT
GTC^TGGAGAAACACATGGAATCCTGTCTCCTGTAGTTGGCTACAGTC^TAAATGAAGGAATGC
CTGAGAGGAAACACTCATATACCTAATGGGAGATCCACGATCCGCCAGTGGATGGAATACTT
CAAAGCCAGTATCAACACACTGAAGTCAGTGGATCTGCATGAATCTGGAAAATTGCAACACC
AGCTTGCTGTGGATTGGATAGGAACATCAACATCCAGTGGCAAAATATTGTTATCCCTG
ATAGGATCAATGACCTATTCACTGTCATTGTTATTCCCTGGGCCAGCATTTCAGACCCCTTCCA
TGGAGGAGAAAAAAATACTGTCATTGTTATTCCCTGGGCCAGCATTTCAGACCCCTTCCA
TTGATGTTTTATCGAAGGGCCCTCAATGTCCACAAAGCCATTCACTGAGATCTTCTGAGA
AGCCCAGACACTATGTTATCATCAAACAGAAAACATCAGGGAGATGTACAATGATGCAGA
AAGATTTAGTGACTTCATGGTTACATTCAATATCTCATCATAAAGGACATTTCAGGATC
TCAGTGTGAGTATCATTGATGCCTGGGATATAACAATTGCATATGGCACAAATAATGTACAC
CCACCTCAACATGTAGTCGGAAATCAGATTAATATTAAACTATTTGTTAAATAACAA

FIGURE 12

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108696
><subunit 1 of 1, 544 aa, 1 stop
><MW: 62263, pI: 9.17, NX(S/T): 7
MKISMINYSKLLALLFILASWIIFTVFQNSTKWSALNLSISLHYWNNSTKSLFPKTPLI
SLKPLTETELRIKEIIIEKLDQQIPPRPFTHVNTTSATHSTATILNPRDTYCRGDQLHIL
LEVRDHLGRRKQYGGDFLRARMSSPALMAGASGKVTDFNNGTYLVSFTLFWEGQVSLSL
LIHPSEGVSALWSARNQGYDRVIFTGQFVNGETSQVHSECGLILNTNAELCQYLDNRDQEG
FYCVRPQHMPCAALTHMYSKNKKVSYLSKQEKSLSFERSNVGVEIMEKFNTISVSKCNKET
VAMKEKCKFGMTSTIPSGHVWRNTWNPVSCSLATVKMKECLRGKLIYLMGDSTIRQWMHEY
FKASINTLKSVDLHESGKLQHQLAVDLDRNINIQWQKCYPLIGSMTYSVKEMEYLTRAI
DRTGGEKNTVIVISLGQHFRPFIDVFIRRNLNVHKAIQHLLLSPDTMVIIKTENIREM
YNDAEFRSDFHGYIQYLIIKDIFQDLSVSIIDAWDITIAYGTNNVHPPQHVVGQINILL
NYIC
```

Important features of the protein:**Signal peptide:**

Amino acids 1-22

N-glycosylation sites:

Amino acids 29-33; 38-42; 47-51; 48-52; 92-96; 160-164; 210-214

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 262-266

Tyrosine kinase phosphorylation site:

Amino acids 236-243; 486-494

N-myristoylation sites:

Amino acids 206-212; 220-226; 310-316; 424-430; 533-539

Amidation site:

Amino acids 127-131

Cell attachment sequence:

Amino acids 113-116

FIGURE 13

GCAAAGAGAAGACTGAAAGACAAACCTGGGTGCAGCCAGAGAGGTCCAGATAGATGAGCTTG
TGGCATCCATTCCCCAAGTTCAGCCTAGGGACTCCACGTACCCAGCTGGGTCTCATTGTTTC
CAGAACTGCATTAGTTAAGATTACCCAGACTGGATTCAAAGGAATACTTCAATTGTTCCG
TCTGTAACACGAAGTAATTGGGCCAGCTGGATGTCAGGATCGTGTGGTTACCATTGTAAT
CTTGCTCTGCTTTGCAAAGCGGCTGAGCTGCGCAAAGCAAGCCCAGGCAGTGTGAGAAGCC
GAGTGAATCATGGCCGGCGGGTGGAGGCCGGAGAGGCTCAAACCGGTCAAACGCTACGCA
CCAGGCCTCCCCTGACGTGTACACATATCTCCATGAGAAATACTTAGATTGTCAGAAAG
AAAATTAGTTATGTCGCTGGTGGCTCAGGATTGCTGCACATGCTGCTAGCAAGAA
ACAAGATCCGCACATTGAAGAACACATGTTCCAAGTTAAAAAGCTGAAAAGCCTGGAT
CTGCAGCAGAATGAGATCTCTAAAATTGAGAGTGAGGCCTTGGTTAAACAAACTCAC
CACCCCTCTTACTGCAGCACAAACCAGATCAAAGTCTTGACGGAGGAAGTGTTCATTACAC
CTCTCTGAGCTACCTGCGTCTTATGACAACCCCTGGCACTGTAATTGAGATAGAAACG
CTTATTCATGTTGAGATTCCAGGAACCGGAATTGGGAACACTACGCCAAGTGTGAAAG
TCCACAAGAACAAAAAAATAAAACTGCGGCAGATAAAATCTGAACAGTTGTAATGAAG
AAAAGGAACAATTGGACCCGAAACCCCAAGTGTCAAGGGAGACCCCCAGTCATCAAGCCTGAG
GTGGACTCAACTTTGCCACAATTATGTTCCATACAAACACTGGACTGCAAAAGGAA
AGAGTTGAAAAAGTCCAAACACATCCCTCCAGATATTGTTAAACTTGACTTGTCA
ATAAAATCAACCAACTCGACCCAAAGGAATTGAAAGATGTTCATGAGCTGAAGAAATTAAAC
CTCAGCAGCAATGGCATTGAATTCATCGATCCTGCCGCTTTTGTAGGGCTCACACATTAGA
AGAATTAGATTATCAAACACAGTCTGCAAAACTTGACTATGGCGTATTAGAAGACTTGT
ATTTTGAAACTCTGGCTCAGAGATAACCCCTGGAGATGTGACTACAACATTCACTAC
CTCTACTACTGGTAAAGCACCCTACAATGTCCATTAAATGGCCTGGAATGCAAAACGCCT
GAAGAATACAAAGGATGGTCTGGAAAATATATTAGAAGTTACTATGAAAGATGCCCAA
AGACAAGTTACCAAGCATATCCTGAGTCATTGACCAAGACAGAAGATGATGAATGGGAA
AAAAACATAGAGATCACACCGCAAAGAAGCAAAGCGTAATAATTACTATAGTAGGATAAGGT
AGAAAATTGTTCTGATTGTAATTAGTTGTATTTCTATACGGTGTAGAAAACATATGTT
TACATTGATTAACTGTGTTGCCTATTATGCAGGGTAATCAGCTAAAGGAAGCTTCTT
AATTATAAGTATTATTGTGACTATTATAGTAATCAAGAGAATGCTATCATCCTGTTGCCTG
TCCATTGTTGGAACAGCATCTGGTGATATGCAATTCCACACTGGTAACCTGCAGCAGTGGG
TCCTAATGATGGCATTAGACTTCAATGTCCTGTATAATGTTTTACTGCTTTAGAAA
ATAAAGAAAAAAACTGGTTCATGTTAAAA

15/246

FIGURE 14

```
>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA117935
><subunit 1 of 1, 440 aa, 1 stop
><MW: 51670, pI: 8.70, NX(S/T): 2
MRVVTTIVILLCFCKAAELRKASPGSVRSRVNHGRAGGGRRGSNPVKRYAPGLPCDVYTYL
HEKYLDQCERKLVYVLPGWPQDLLHMLLARNKIRTLKNNMFSKFKLKSDLQQNEISKI
ESEAFFGLNKLTTLLQHNQIKVLTTEEVFIYTPLSYLRLYDNPWHTCEIETLISMLQI
PRNRNLGNYAKCESPQEQQNKLRQIKSEQLCNEEKEQLDPKPQVSGRPPVIKPEVDSTF
CHNYVFPIQTLDCRKELKKVPNNIPPDIVKLDLSYNKINQLRPKEFEDVHELKKLNLSS
NGIEFIDPAAFLGLTHLEELDSNNSLQNFDYGVLEDLYFLKLLWLRDNPWRCDYNIHYL
YYWLKHMHYNVHFNGLECKTPEEYKGWSVGKYIRSYEYECPKDKLPAYPESFDQDTEDEW
EKKHRDHTAKKQSVIITIVG
```

Important features of the protein:**Signal peptide:**

Amino acids 1-15

N-glycosylation sites:

Amino acids 297-301; 324-328

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

Amino acids 19-23; 39-43; 430-434

N-myristoylation sites:

Amino acids 24-30; 37-43

Amidation site:

Amino acids 37-41

16/246

FIGURE 15

GGGGCAGCAGCGCGGGCCCCAGCAGCCTCGGCAGCCACAGCCGCTGCAGCCGGGCAGCCTC
CGCTGCTGTCGCCTCCTCTGATGCGCTTGCCTCTCCCAGGGCCGGACTCCGGGAGAATGT
GGGTCTAGGCATCGCGCAACTTTGCGGATTGTTCTGCTTCCAGGCTTGCGCTGCAA
ATCCAGTGCCTACCAGTGTGAAGAATTCCAGCTGAACAAACGACTGCTCCTCCCCGAGTCAT
TGTGAATTGCACGGTGAACGTTCAAGACATGTGTCAGAAAGAAGTGTAGGGAGCAAAGTGC
GGATCATGTACCGCAAGTCTGTGCATCATCAGCGGCCTGCTCATGCCCTTGCCGGTAC
CAGTCCTTCTGCTCCCCAGGGAAACTGAACCTCAGTTGCATCAGCTGCTGCAACACCCCTCT
TTGTAACGGCCAAGGCCAAGAAAAGGGGAAGTTCTGCCTCGGCCCTCAGGCCAGGGCTCC
GCACCAACCATTGTCCTCAAATTAGCCCTTCTCGGCACACTGCTTGAAGCTGAAGGGAGA
TGCCACCCCTCCTGCATTGTTCTCAGCCCTGCCCAACCCCCACCTCCCTGAGTGA
GTTTCTTCTGGGTGTCCTTTATTCTGGTAGGGAGCGGGAGTCCGTGTTCTCTTTGTTCC
TGTGCAAATAATGAAAGAGCTCGGTAAAGCATTCTGAATAAATTCAGCCTGACTGAATTTC
AGTATGTACTTGAAGGAAGGGAGGTGGAGTGAAAGTTCAACCCCATGTCGTGTAACCGGAGT
CAAGGCCAGGCTGGCAGAGTCAGTCCTAGAAGTCACTGAGGTGGCATCTGCCCTTGTAA
AGCCTCCAGTGTCCATTCCATCCCTGATGGGGCATAGTTGAGACTGCAGAGTGAGAGTGA
CGTTTCTTAGGGCTGGAGGGCCAGTTCCACTCAAGGCTCCCTGCTTGACATTCAAACCTT
CATGCTCCTGAAAACATTCTGCAGCAGAATTGGCTGGTTCGCCCTGAGTTGGCTCT
AGTGAAGACTCAATGACTGGACTTAGACTGGGCTCGGCCTCGCTCTGAAAAGTGT
TAAGAAAATCTCTCAGTTCTCCTTGCAAGAGACTGGGCCGGACCGAAGAGCAACGGGC
GCTGCACAAAGCGGGCGCTGTCGGTGGAGTGCAGTACGCGCAGGCCTCTCGTGG
TTGGCGTGCAGCGACAGGCGGCAGCACAGCACCTGCACGAACACCCGCCAAACTGCTG
CGAGGACACCCTGTAAGGAGCGGGTTGATGACCGAGCTGAGGTAGAAAAGTCTCCGAGA
AGGGGAGGAGGATCATGTACGCCCGAAGTAGGACCTCGTCCAGTCGTGCTTGGTTGGCC
GCAGCCATGATCCTCGAATCTGGTTGGCATCCAGCATACGCCAATGTCACAACAATCAG
CCCTGGCAGACACGAGCAGGAGGGAGAGACAGAGA

17/246

FIGURE 16

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA119474
><subunit 1 of 2, 141 aa, 1 stop
><MW: 15240, pI: 8.47, NX(S/T): 1
MWVLGIAATFCGLFLLPGFALQIQCYQCEEFLQLNNDCSSPEFIVNCTVNVQDMCQKEVME
QSAGIMYRKSCASSAACLIASAGYQSFCSPGKLNSVCISCCNTPLCNGPRPKRGSSASA
LRPGLRTTILFLKLALFSAHC
```

Important features of the protein:

Signal peptide:

Amino acids 1-22

N-glycosylation site:

Amino acids 45-49

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 113-117

N-myristoylation sites:

Amino acids 5-11;115-121;124-130

Ly-6 / u-PAR domain proteins:

Amino acids 94-107

18/246

FIGURE 17

CGCAAAGCCGCCCTGGGGCGCTC**ATGGCGGGACGCCCTGGGAAAGGCTTAGCCGCGGT**
GTCTCTCTCTGGCCTTGGCCTCTGTGACTATCAGGTCTCGCGCTGCCGCGCATCCAGG
CGTTCAAAACTCGTTCATCTTCTTGGTTCATCTTAATACCAACGTCTGGTTCT
AATGGTTCCAAAGAAAATTCTCACAATAAGGCTGGACGTCTCCTTACCCAGGTCAAAAGT
TGAACGAAGCCAGGTTCTAATGAGAAAGTGGGCTGGCTTGTGAGTGGCAAGACTATAAGC
CTGTGGAATACACTGCAGTCTGTCTGGCTGGACCCAGGTGGGAGATCCTCAGATCAGT
GAAAGTAATTCTCCCAGTTAACGAAAAGGATGGGATGTTGAGGAGAAAGAGCAAGAA
TGGCCTGTATGAGATTGAAAATGGAAGACCGAGAAATCCTGCAGGACGGACTGGACTGGTGG
GCCGGGGCTTGGGCGATGGGCCAAATCACGCTGCAGATCCCATTATAACCAGATGG
AAAAGGGATAGCAGTGGAAATAAAATCATGCATCCTGTTCTGGGAAGCATACTTACAATT
TGTGCAATAAAAGGAAAGACTGTGGAGAATGGCAATCCCAGGGGGATGGTGGATCCAGGA
GAGAAGATTAGTGCACACTGAAAAGAGAAATTGGTGGAGGAAGCTCTCAACTCCTTACAGAA
AACCACTGCTGAGAAGAGAGAAATAGAGGAAAAGTTGCACAAACTCTCAGCCAAGACCACC
TAGTGATATATAAGGGATATGTTGATGATCCTCGAAACACTGATAATGCATGGATGGAGACA
GAAGCTGTGAACCTACCATGACGAAACAGGTGAGATAATGGATAATCTTATGCTAGAAGCTGG
AGATGATGCTGGAAAAGTGAATGGTGGACATCAATGATAAACTGAAGCTTATGCCAGTC
ACTCTCAATTCAAACTTGCTGAGAAACGAGATGCACACTGGAGCGAGGACTCTGAA
GCTGACTGCCATGCCTG**TAG**CTGATGGTCTCCGTGTAAGCCAAGGCCACAGAGGAGCAT
ATACTGAAAAGAAGGAGTACACAGAAATTATACTATAAAAGGGCAGGGTAGGCCACTTG
GCCTATTACTTCAAAACAATTGCATTAGAGTGGTTCGCATCAGAATAACATGAGTAAG
ATGAACCTGGAACACAAAATTTCAGCTTTGGTCAAAAGGAATATAAGTAATCATATTG
TATGTATTGATTAAGCATGGCTAAATTAAACAAACTAATGCTTTGAAGAAC
ATAATCAGAATAAGATAATTCTGATCAGCTATA

19/246

FIGURE 18

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA119498
><subunit 1 of 1, 350 aa, 1 stop
><MW: 39125, pI: 8.53, NX(S/T): 2
MAGRLLGKALAAVSLSLALASVTIRSSRCRGIQAFRNSFSSWFHLNTNVMSGNSKEN
SHNKARTSPYPGSKVERSQVPNEKVGWLVEQDYKPVEYTAWSLAGPRWADPQISESNF
SPKFNEKDGHVERKSKNGLYEIENGRPRNPAGRTGLVGRGILLGRWGPNAADPIITRWKR
DSSGNKIMHPVSGKHLQFVAIKRKDCGEWAIPGGMVDPGEKISATLKREFGEALNSLQ
KTSAEKREIEEKLHKLFSQDHLYIYKGYVDDPRNTDNAWMETEAVNYHDETGEIMDNLML
EAGDDAGKVKWVDINDKLKLYASHSQFIKLVAEKRDAHWSEADCHAL
```

Important features of the protein:**Signal peptide:**

Amino acids 1-20

N-glycosylation site:

Amino acids 55-59

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 179-183

N-myristoylation sites:

Amino acids 53-59; 56-62

mutT domain signature:

Amino acids 215-235

20/246

FIGURE 19

CGAGGGCTCCTGCTGGTACTGTGTTCGCTGCTGCACAGCAAGGCCCTGCCACCCACCTTCAG
GCCATGCAGCCATGTTCCGGGAGCCCTAATTGCACAGAAGCCCATGGGGAGCTCCAGACTGG
CAGCCCTGCTCCTGCCCTCCTCCCTCATAGTCATCGACCTCTGACTCTGCTGGGATTGGC
TTTCGCCACCTGCCCTACTGGAACACCCGCTGTCCTCTGCCCTCCCACACGGATGACAGTTT
CACTGGAAGTTCTGCCTATATCCCTGCCGCACCTGGTGGGCCCTCTTCTCCACAAAGCCTT
GGTGTGTGCGAGTCTGGCACTGTTCCCGCTGTTGTGCCAGCATCTGCTGTCAAGGTGGCTCA
GGTCTTCAACGGGGCTCTTCCACCTCCTGGTGCAGAAATCaaaaAGTCTTCCACATTCAA
GTTCTATAGGAGACACAAGATGCCAGCACCTGCTCAGAGGAAGCTGCTGCCTCGTCACC
TGTCTGAGAAGAGCCATCACATTCCATCCCTCCCCAGACATCTCCCACAAGGGACTTCGC
TCTAAAAGGACCCAACCTCGGATCCAGAGACATGGGAAAGTCTTCCAGATTGGACTCACA
AAGGCATGGAGGACCCGAGTTCTCCTTGATTGCTGCCTGAGGCCGGGCTATTGGGTGA
CCATATCTCAGGCCCTGAGGTCAAGCGTGCCTTGTCACCAGTGGCACTGGAGTGTGAA
GAGCTGAGCAGTCCCTATGATGTCAGAAAATTGTTGTCTGGGGCCACACTGTAGAGCTGCC
TTATGAATTCTTCTGCCCTGTCGTGCTAGAGGCATCTACCTGCAAGAGGACACTGTGA
GGCGAAAAAAATGTCCTTCCAGAGCTGGCCAGAACGCTATGGCTCGGACTCTGGAAAGTCA
GTGCACTTCACTGACTACAGCCAGCACACTCAGATGGTATGCCCTGACACTCCGCTGCC
ACTGAAGCTGGAAGCTGCCCTGCCAGAGGCACGACTGGCATAACCTTGCAAAAGACCTCC
CGAATGCCACGGCTCGAGAGTCAGATGGGTGGTATGTTTGAGAAGGGTGGACCTGCACCCC
CAGCTCTGCTCAAGTTCTTTGGAAACAGCAGCCATGTTGAATGCCCCACCAGACTGG
GTCTCTCACATCCTGGAATGTAAGCATGGATACCCAGGCCAGCAGCTGATTCTCACTTCT
CCTCAAGAATGCATGCCACCTTCAGTGCTGCCCTGGAGCCTCCAGGCTGGGGCAGGACACT
TTGGTGCCCCCGTGTACACTGTCAGCCAGGCCGGGCTCAAGCCCAGTGTCACTAGACCT
CATCATTCCCTCCTGAGGCCAGGGTGTGCTGTCCTGGTGTGGCGGTCAAGATGTCAGTTG
CCTGGAAGCACCTTTGTGTCAGATGTCCTTACAGACACCTGGGCTCTTGATCCTGGCA
CTGCTGCCCTCCTCACCCACTGGGTGTGCTGTCCTGCCCTGCACGCCGGACTCGGAGGCCAGCGGC
AGGCCCGGGCCAGCGCGCCAGTGCTACGGCAGCGCTGGCGGCCGGCGACGTGATC
GCCTGGTGGAGCGCTGGCTGAACGTCTACGGCAGCGCTGGCGGCCGGCGACGTGATC
GTGGACCTGTGGAGGGGAGGCACGTGGCGCGTGGGCCGCTGCCGTGGCTCTGGCGGC
GCGGACGCGCGTAGCGGGAGCAGGGCACTGTGCTGCTGCTGGAGCGGCCGACCTTC
GCCCGGTAGCGGCCCGACCCCGGCCGCGCCCTGCTGCCCTGCTCCACGCTGCCCG
CGCCCGCTGCTGCTGCTGCTTACTTCAGTCGCCTCTGCGCCAAGGGCGACATCCCCCGCC
GCTGCGGCCCTGCCGCTACCGCCTGCTGCCGACCTGCCGTCTGCTGCCGCTGG
ACCGCGGCCCTTCGAGAGGCCACCAGCTGGGCCCTGGGGCGCCAGCGCAGGCAG
AGCCGCCTAGAGCTGTCAGCCGGCTGAACGAGAGGCCGCCACTTGCAAGACCTAGGTTG
AGCAGAGCTCCACCGCAGTCCGGGTGTCT

21/246

FIGURE 20

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA119502
><subunit 1 of 1, 667 aa, 1 stop
><MW: 74810, pI: 9.55, NX(S/T): 3
MGSSRLAALLLPLLIVIDLSDSAGIGFRHLPHWNTRCPLASHTDDSFHGSSAYIPCRTW
WALFSTKPWCVRVWHCSRCLCQHLLSGGGSLQRGLFHLLVQKSKSSTFKFYRRHKMPAP
AQRKLLPRRHLSEKSHHISIPSPDISHKGLRSKRTQPSDPETWESLPRLDSQRHGGPEFS
FDLLPEARAIRVTISSLGPEVSVRLCHQWALECEELSSPYDVQKIVSGGHTVELPYEFLLP
CLCIEASYLQEDTVRKKCPFQSWPEAYGSDFWKSVHFTDYSQHTQMVMALTLRCPLKLE
AALCQRHDWHTLCKDLPNATARESDGWYVLEKVDLHPQLCFKFSFGNSSHVECPHQGTGSL
TSWNVSMDTQAQQLILHFSSRMHATFSAAWSLPGILGQDTLVPPVYTVSQARGSSPVSDL
IIPFLRPGCCVVLWRSVDQFAWKHLLCPDVSYRHLGLLILALLALLTLLGVVLALTCCR
QSGPGPARPVLLLHAADSEAQRRLVGALAEELLRAALGGGRDVIVDLWEGRVARVGPLPW
LWAARTRVAREQGTVLLLWSGADLRPVSGPDPRAPLLALLHAAPRPLLAYFSRLCAK
GDIPPLRALPRYRLLRDLPRLLRALDARPFAEATSWGRLGARQRQSRLELCRLESEA
ARLADLG
```

Important features of the protein:**Signal peptide:**

Amino acids 1-23

Transmembrane domain:

Amino acids 455-472

N-glycosylation sites:

Amino acids 318-322; 347-351; 364-368

Glycosaminoglycan attachment site:

Amino acids 482-486

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

Amino acids 104-108; 645-649

Tyrosine kinase phosphorylation site:

Amino acids 322-329

N-myristoylation sites:

Amino acids 90-96; 358-364; 470-476

Eukaryotic cobalamin-binding proteins:

Amino acids 453-462

FIGURE 21

CGGCTCGAGGCCCTTGTGAGGGCTGTGAGCTGCCCTGACGGTGGCACCA**ATGAGCAGCTCA**
GGTGGGGCGCCCGGGCGTCCGCCAGCTCTGCCGCCGCCGCAGGAAGAGGGCATGACGTG
GTGGTACCGCTGGCTGTGCGCTGTCTGGGTGCTGGGGCAGTCTTGCAGCTCTG
GCCTCTTCAACTGCATCACCATCCACCCCTGAACATCGCGGCCGGCTGTGGATGATCATG
AATGCCTCATCTTGTGCTGTGAGGGGCCCTCTGCTGCCAGTTCATCGAGTTGCAAA
CACAGTGGCGGAGAAGGTGGACCGCTGCCCTGGCAGAAGGCTGTCTTACTGCC
TGGCGGTGCTTCCCACATCGTACAGCCTGACCCCTGACCACGCTGCTGGCAACGCCATGCC
TTGCTACGGGGTGTGTACGGACTCTGTCTGGCAAAAGGGCGATGCGATCTCTA
TGCCAGGATCCAGCAGCAGAGGCAGCAGGGATGAGGAGAAGCTGCCAGACCCCTGGAGG
GGGAGCTG**TGA**AGGGCTGGGCCCTCCCTCCCTGTCCCCCTTCTGGCTCTGTGGTC
CAAGTGGGACTGGACTGTCCACGCTGAGGCACAGCCTGGAGAGGGGCCCTTGACGTGTCC
CTACACCTGGAGTCCTCTGCTCCTTCTCCAGACTGGCTTAAGCCAGGAGCCACTGGCTGCT
GGTGTAGGGTCTGGCTGTTGACTTGAGGCAGAGCCTGCCAGCAGCTGTGTGGACACTACC
CAGCCCTACTCCTCTGCTGGTGGCTGCAGATCTCACACCACAGACAGGGCTGCCGTGTA
CCTGCTGTGACCTGGAGCAGCTCCCCCTGGAGATGCTGGCTGGCTTGAGGGAGGGCA
AGTGGGACCCCTGCCACCTGGCACTGAGCAGAGGACCTCCCCAGCTCTTAGCAGGTGG
AGCCCCAGGGCTGGACAGCCTGCCAGCAACCTCCACTGCTGCCCTAGGGTGCAG
CGCCCACTGTCACCCCTGCCCTCTGAAGAAGCCCACAGGGCTCTAAGGTGCACCCCGGTACC
TGAAGACTGCAAGCCTGGCAGTGACTGGACAGCTGGTGGGGATGCTCCCTGCTGGCCCTGG
GAACCTGGACAGGCCACCTCAAGGCCCTCGGCTGCCCTCCCTGGCCTGCTGGGC
CCCTAGGTTCTACCCATCACCCCCCGCCCTGCTGGCCTGGTCTAAGGAAGTGGGAGAG
CAGGCTCTCCCTGGCACCGAGGGTGCCACCCCTCCCTGGTGTGGCCCCGTCAACATCAGC
CACAGCCCAGCCCCATTAGTGGGTTAGTGGGTCTGACCTCAGCCCCACTCAGGTGCTCCTGC
TGGCCTGCCCAAGCCCTGCCCTCAGGGAGCTTCTGCCCTTAAGAACTGGGAGAGGCCACAGT
CACCTCCCCACACAGAGCTGCCCCACTGCCCTGGTGCACCCCTGCTGCCCTGT
CCCAGGGAGGATGCAAGAGAGCTGGTGCCAGGATGTGCACCCCTATTCCCTGCTGCCCTGT
GCCCTAGCCCGCTGCCCTCTGACCGTGAGGCTGGCTCTGCCATCAGGCCATGGCAGGTGCTG
GTCAGGCCTGGCTTAGCCCAGGTGGGCTTGGCAGAAGCAGGGGGGTGTGGAAGATATTCCA
TCTGGGCCAACCCCAGGCTGGGCTGCCCTGCGCTGAGCTCTGGAGCGCAGGTACTGGGCTTGC
TAAGTGAACTGTTCCCAGGAACACCTCTGGGCCATCTGCGCTGAGGCTGGGAGTGGCA
TCTGAGGCCGGAGTGGCATCTGAGGCCAGGAGTGGCAGGGCTGGTGGCTGGCTGGCGTGGGTT
TTCTGGCCCTGCCAGTACTGCCCTGGGACTTGGTGGCTCTGGCTCAGCAGCATCCCA
CCCCTGGAGTCTGGCAGCTGAGCCCAGGGTGGCAGGGCATTATAGCCTGGTGGACATG
TGCCTTCAGGGTCCCTGGGCCACCTCCTCAGGCCAGTGCTGGGTTCAAAGGGCTGTGT
GTGTGTGTGTTGTGTGTATGTATATGTGTGTTGGTGCACACATCTGCTCCATGTATGCA
GTGAGACCTGTCACCTCCCACAAGGAGCAAGGGCTCTGCCGCCCTTGCTCATCCCTACC
CAGGTAGTGGGACCCGGGCCCCCTTGCCCTGGCTGCCCTGCTCTGCCCTTCCAGAGGG
GTCTCACTGACAGCCAGAGACAGCAGGAGAAGGGTTGGCTGGATCAAGGAAGGCTGCC
TGTACCCCTGTGGGAAATGGTGGTGCATGGCTGGATGCAGAGGTGGAAGGCCCTGGCAC
AGGCAGAGTGGCGTGTACCTGCTCCAGGTTCCAGCAAGTCTGCAAGCTGCAAGTCC
GGTCCCTGACCCCTGCGCCAGGGGGCGTGTGCTGCCAGCAGGGGCCCTGCCCTGCAAGGAA
CGTCTCTCCGGCGGCTGGGCCGCTCTGCCCTGGTGTGGCTGTGTGGCGCCCTTCC
TTGTTGTTCCCTGTGTTCTGTGCTTAAGCAATAAGCGTGGCGTGGAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAA

23/246

FIGURE 22

```
>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA119516
><subunit 1 of 1, 172 aa, 1 stop
><MW: 18470, pI: 5.45, NX(S/T): 0
MSSGGAPGASASSAPPAQEEGMTWWYRWLCRLSGVLGAVSCAISGLFNCITIHPNIAA
GVWMIMNAFILLLCEAPFCCQFIEFANTVAEKVDRRLRSWQKAVFYCGMAVPIVISLTLT
TLLGNAIAFATGVLYGLSALGKKDAISYARIQQQRQQADEEKLAETLEGEL
```

Important features of the protein:**Signal peptide:**

Amino acids 1-42

Transmembrane domains:

Amino acids 64-77;109-128

Tyrosine kinase phosphorylation site:

Amino acids 142-150

N-myristoylation sites:

Amino acids 5-11;6-12;9-15;35-41;38-44;46-52;124-130;132-138

Amidation site:

Amino acids 140-144

FIGURE 23

GTGAAACACCCATGGTTTATGCTCTATTCTCTTCCATCTTCCACATCCTCTT
CTGAATGTATCAAACACTTCCTTGAAGTGGGGCACCAGGAGGGCCTCCAGTCTCCAATG
CAGGGACTCAGGGCAGGGATCTCTGAGAAAGTGGCCATCTCGTTATTAAAGCTCTGTCCCTC
TGCTTCCCTCTCACCTCAGAACAGCCGTTATTCAACAGAGCTCCAGGTTGCCAGCTAGG
GGTTTCGGGACCATAGACCAAGCAACCCGAGAGACTGAGTACTGACCTGCAGTTGTTCCAG
AAACTCTGCTGGATTAGGTTGTGACCTAGAAGTGAACTTGAACACTAACAGTGAGAAGGCAG
GGTAAGAATGCAGTCTAGAGCGAACCTTCTCCACTAGACTTGTAAGTAATTAAAGTGAAT
CCTGTCCCCCTGGGTTCTATCCTGGCTGGCTCTGCTGGTGAACTTGACTGGCCAGCATAGG
GCACTTGATGAGACCCCTGGAATGCTGAGGCCAGTTGGCAGCAAGCTTCACCTCATCCTTC
TGCCCCATCTATCCAGCCATTCAAACATTCTGCCTGAAGACATTATCAAGCTCCTGCAA
TGGGTCAAGGCATCTGCTAGGCACTGGGACACAGAGCTCACAGTCTCCTGGAGGGGGTGAGA
GATGACTGACAGGTGGCTGTGGTGCAGTGTGACCTGGGAATGCACACAGTACTGTGGAAAC
ACGGGAGAGGCATCTAGCACAACTGAGAGGCCAGGGGAGGCTTCCTGGCAGGTTCCCT
TAACCATCTAAGGGAAAGAGGCCTAGGTAGGAAAATAAGGGACAGTGGTGTCCCAGACA
GAGGGCACTTACATGGAA

25/246

FIGURE 24

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA119530
><subunit 1 of 1, 113 aa, 1 stop
><MW: 12799, pI: 7.53, NX(S/T): 1
MVLCSISLFLIFFHILFLNVSNYFLEVGHQEGHSSLQCRDSGAGISEKVAISLLKLCPLL
PSHLRSSPFIQQSSRLPARGFRDHRPSNPERLSTDQLFQKLCWELGCDLEVN

Important features of the protein:

Signal peptide:

Amino acids 1-18

N-glycosylation site:

Amino acids 19-23

Glycosaminoglycan attachment site:

Amino acids 41-45

N-myristoylation site:

Amino acids 42-48

26/246

FIGURE 25

CGGGCTCCGCGCGGTCCCACCTCCGGCTCCCTCGCCTCCAGGATGCGCTGAGCCCTACAA
CACCCCCAGCGGCCGCCGGCTCCCCCAGAGGTGTGAATGACAGAGGTGGTGCCATCCAGCG
CGCTCAGCGAGGTCAAGCCTGCCCTCGCCACGATGACATAGACACTGTGAAGCACCTG
TGTGGCGACTGGTCCCCATCGAGTACCCAGACTCATGGTATCGTGAATCACATCCAACAA
GAAGTTCTTCCCTGCTGCAACCTACAGAGGTGCCATTGTGGGAATGATAGTAGCTGAAA
TTAAGAACAGGACAAAATACATAAGAGGATGGAGATATTCTAGCAACCAACTCTCTGTT
GACACACAAGTCGCGTACATCTAACAGTCTGGCGTGTGAAAGAGTTAGGAAGCACGGCAT
AGGTTCCCTCTTACTGAAAGTTAAAGGATCACATATCAACCACCAGGACACTGCA
AAGCCATTACCTGCATGTCCCTACCACCAACAACACAGCAATAAACACTTCTATGAAAACAGA
GACTTCAAGCAGCACCCTATCTCCCTATTACTACTCCATTGAGGGGTCTCAAAGATGG
CTTCACCTATGTCCTCTACATCAACGGGGGACCCCTCCCTGGACGATTGGACTACATCC
AGCACCTGGGCTCTGCACTAGCCAGCCTGAGCCCTGCTCCATTCCGCACAGAGTCTACCGC
CAGGCCACAGCCTGCTGCACTGGCATGGTGGGCATCTCTCCAAGAGTGGCAT
CGAGTACAGCCGGACCATGTGATGTCGGCTGGCAGCCGACCCAGGGCCACCCCTCAGCC
GCCCGCAGAGCCGGCTTCCTGTCCATCTGACCCCTCTGTTTCTGCAAGGAGCTGCCAGC
CATCTAACTGGGCTCGTCGGCTGCCAGCTGCAGGCCGGTGTACACGGGCTGGGAAC
AGAACATCGTGGCATGCGCAGAGCATGCCATCCGTGGCAGGCTCTCAGCTCCCTCCCT
GCTTCTGAAACCTCTGCCTGCTGCCCTGGCCCTGCCCCCTGCGCATGCACCATCCCCAGG
GCTGACCCAGTGTGGCTGCATTCACTGGGAGGGGCTGCCCTCACTGGGCTCTCCACTCCG
CTGCCTGTTCTGCACTCCTGGAAAGCTGGAGGGGACTTCTCTGCAAGGGAGGAA
CGCAAGTATTATGGACACACTTGACCGTAAAGGCACAGGAGCCTCGGAACAAGGGCGCAA
TAAAGGAAATGGCCCGTCCCCTCCAGAACCAAGAGCCTGGGGGTGAGGAGTGG
CCCCCACTCCTCCATGAGGGCTGATGAGGGGTGGCAGCCTGGGGAGGCTTCCTCGCAA
GCACAGAGCTCTGAGGCTCAGCCCCCTGGCACAGGGCGTCAGCATTGAGGACGTTCTACT
CCTCAGCACCTCCGTGCAGTTACAGTCCCTGGGAGGTACACTGCCGTCGGACCTTGG
CATGCTCCATTCACTGACCTGCTGAGGACAGGCATGCCGAGACTCCTGGGCTCCCCG
CCCTCCCTCATGCTGCCACAAGCTGCTCCAAGGCCATGCAAGACAGGAGGAAAG
CTGAGCTGACATTAGGCCTCAAGGCTGCCATCTGTCTTGTAGGGCCTGGCCTTGTGGCAG
GGGGCAGTCCTGTGCCTTGTGGGCCCTCAGCCTCTGAGGGCAGAGATGCTGTCAGTGC
GGTGCATCACATACTCTAGCATCCTCTCCACCCCTGCATTCAAATGCTGCTGCTGCCTGC
CCTGCCCTCCGATGCAGGGGTGGGTGGGGGGCGGAGTCCCAGCAGCATAGCTGCAGTGT
ACAAAGCCATGGCAGAGGTCTAGCGGCCACCCCTGCCAGCCTGAGGAGGAGGAGAG
GGAGGAACAAACCCCTGGCAGACGGGTCTCAGGGACCTGTGTCCTTCCGCTCCAGAGCTGC
CCAGCCACGGCTCTCAGGGTGTGGGCAGCCCCAGGTCCCTTGAACCTCAGCTGGGGC
CAGGGGCCCTCAGAATGAAGGCAGGCACCAGGCAGGAGCAGCATCCCCCTCCTGACGGTGC
TGGCAGGAGGGCCGCCATGCTGACTGCTGAACCTCTGCTGACCTGACAGTGCTGGCGGG
AGGGCCGCACCATGCTGACTGCTGAATCTCTGCTGAGGCTGCCTGCCTGCCGGCCAGCT
CAGGCCCTCTCCACTGCAATCAGTGGCAGTCATGTGATTCTATTCTGCCACAGGGT
AAGGGACGAGTCTCTGGAAGGCTCTGCCATGGACATTGTCTCGGGCTCAGAGGCCAC
CCTGCCACACCTGCCCTAATCACTGCACTGCAAGTGTGAACAGATTGTAGCG
TTCTGTCTCATTACGAGCAAATAAGACTTCAATTGGAAAAA

27/246

FIGURE 26

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA121772
><subunit 1 of 1, 242 aa, 1 stop
><MW: 27465, pI: 7.72, NX(S/T): 3
MTEVVPPSSALSEVSLRLLCDDIDTVKHLCDWFPIEYPDSWYRDITSNKFFSLAATYR
GAIVGMIVAEIKNRTKIHEDGDILATNFSVDTQVAYILSLGVVKEFRKHGIGSLLLES
KDHISTTAQDHCKAIYLHVLTNNNTAINFYENRDFKQHHYLPYYY SIRGVLKDGFTYVLY
INGGHPPWTILDYIQHLGSALASLSPCSI PHRVYRQAHSLCSFLPWSGISSKSGIEYSR
TM

N-glycosylation sites:

Amino acids 73-77; 88-92; 143-147

N-myristoylation sites:

Amino acids 61-67; 65-71; 198-204; 235-241

Matrixins cysteine switch motif:

Amino acids 18-31

28/246

FIGURE 27

GTTGGGCAGCAGCCACCCGCTCACCTCCATCCCCAGGACTTAGAGGGACGCAGGGCGTTGGG
AACAGAGGACACTCCAGGCCTGACCCCTGGGAGGCCAGGACCAGGGC~~AA~~AGTCCCCTGGC
AAGAGGAGTCCTCAGAGGTCTTCATTCA~~G~~CGGTTCCGGGAGGTCTGGGAAGCCCACGGCCT
GGCTGGGGCAGGGTCAACGCCGCCAGGCC~~A~~TGTCCTGTGCTGGCTGCTGCTCTGGTG
ATGGCTCTGCCCTCAGGCACGACGGCGTCAAGGACTGCGTCTTCTGTGAGCTACCGACTC
CATGCAGTGTCTGGTACCTACATGCAC~~T~~GTGGCGATGACGAGGACTGCTTCACAGGCCACG
GGGTGCCCCGGGACTGGTCCGGTCATCAACAAAGGCTGCCTGCGAGCCACCAGCTGC~~G~~
CTTGAGGAACCCGTCA~~G~~CTACAGGGCGTCA~~C~~CTACAGCCTCACCACCAACTGCTGCACC~~G~~
CCGCCTGTGTAACAGAGCCCCGAGCAGCCAGACAGTGGGGCCACCACCGCTGGCACTGG
GGCTGGGTATGCTGCTTCCTCACGTTGCTG~~T~~GACCAACAGGGAGGACAGGGCCTGGGACT
GTTCTCCCAGATCCGCCACTCCCCATGTCCCCATGTCTTCCCCACTAAATGGCCAGAGAG
GCCCTGGACAACCTCTTGC~~G~~CCCTGGCTTCATCCCTCTAAGGCTGTCCACCAGGAGCCCG
GTGCTAGGGGAAGCATCCCCAGGCCTGACTGAGC~~G~~GGCAGGGAGCAGGGCCGTGGGTTGA
TTGTATTACTCTGTTCACTGGTTCTAAGACGCAGAGCTTCTCACATCTCAATCAGGATGCT
TCTCTCCATTGGTAGC~~A~~TTAGAGTCCATGAAATATGGTAAAAAATATATATATCATAA
TAAATGACAGCTGATGTTATGGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

29/246

FIGURE 28

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA125148
><subunit 1 of 1, 124 aa, 1 stop
><MW: 13004, pI: 5.70, NX(S/T): 0
MVLCWLLLLVMALPPGTTGVKDCVFCELTDSMQCPGTYMHCGDDECFTGHGVAPGTGPV
INKGCLRATSCGLEEPVSYRGVTYSLTTNCCTGRLCNRAPSSQTVGATTSLALGLGMLLP
PRL
```

Important features of the protein:

Signal peptide:

Amino acids 1-13

N-myristoylation sites:

Amino acids 19-25;52-58;64-70;81-87;106-112

Ly-6 / u-PAR domain proteins:

Amino acids 84-97

30/246

FIGURE 29

GGCATTGAAAGCCCAGTGTGCCAGGGGCATCTCCTTGTGTTATGAGAGACCTGCA
TTCTCCCTGGCTCAGTTCTCTCAGGCTCTCCAGAGCTCAGGACCTCTGAGAAGAATGGAGCC
CTCCTGGCTTCAGGAACATGGCTCACCCCTTCTGCTGCTGATCCTCCTCTGCATGTCTC
TGCTGCTGTTCAGGAACATGGCTGAGGAGATGGATGATCAGAGCCCTGCAC
CTGTTCTGCACCCCTGCCACTGGTCTATGCCACAAGGAGTTTACCCAGTAAAGGA
GTTGAGGTTGATCATAAGCTGATGGAAAAATACCCATGTGCTGTTCCCTGTGGGTTGGAC
CCTTACGATGTTCTCAGTGTCCATGACCCAGACTATGCCAAGATTCTCCTGAAAAGACAA
GATCCAAAAGTGTGTTAGCCACAAAATCCTGAATCCTGGGTTGGTCAGGACTTGTGAC
CCTGGATGGTCTAAATGGAAAAGCACCGCCAGATTGTGAAACCTGGCTCAACATCAGCA
TTCTGAAAATATTCAACCATGATGTCAGAGTGTGTCGGATGATGCTGAACAAATGGGAG
GAACACATTGCCAAAACACTCACGTCTGGAGCTTTCAACATGTCTCCCTGATGACCTGGA
CAGCATCATGAAGTGTGCCCTCAGCCACCAGGGCAGCATCCAGTTGGACAGTACCCCTGGACT
CATACCTGAAAGCAGTGTCAACCTTAGAAAATCTCAACCAACCGCGCATGAACAAATTTCTA
CATCACAACGACCTGGTTCAAATTCAAGCTCTCAAGGCCAATCTTCTAAATTAACCA
AGAACTTCATCAGTTACAGAGAAAAGTAATCCAGGACCGGAAGGAGTCTCTAAGGATAAGC
TAAAACAAGATACTACTCAGAAAAGGCGCTGGGATTTCTGGACATACTTTGAGTGCCAAA
AGCGAAAACACCAAAGATTCTGAAAGCAGATCTCAGGCTGAAGTAAAACGTTCATGTT
TGCAGGACATGACACCACATCCAGTGTCTATCTCCTGGATCCTTACTGCTTGGCAAAGTACC
CTGAGCATCAGCAGAGATGCCAGATGAAATCAGGAACTCTTAGGGGATGGGTCTTCTATT
ACCTGGGAACACCTGAGCCAGATGCCCTACACCACGATGTGCATCAAGGAATGCCCTCGCCT
CTACGCACCGGTAGTAAACATATCCGGTTACTCGACAAACCCATCACCTTCCAGATGGAC
GCTCCTTACCTGCAGGAATAACTGTGTTATCAATATTGGCTTCTCACCAACACCCCTAT
TTCTGGGAAGACCTCAGGTCTTAAACCCCTGAGATTCTCCAGGAAAATTCTGAAAAAAT
ACATCCCTATGCCCTCATACCATCTCAGCTGGATTAAGGAACCTGCATTGGCAGCATTG
CCATAATTGAGTGTAAAGTGGCAGTGGCATTAACCTGCTCCGCTTCAAGCTGGCTCCAGAC
CACTCAAGGCCCTCCCCAGCCTGTCGTCAAGTTGTCCTCAAGTCCAAGAATGGAATCCATGT
GTTGCAAAAAAAGTTGCTAATTTAAGCCTTCTGTATAAGAATTAAATGAGACAATTTCT
ACCAAAGGAAGAACAAAGGATAATATAACAAAATATGTATATGGTTGTTGACAAA
TTATATAACTTAGGATACTTCTGACTGGTTTGACATCCATTAACAGTAATTAAATTCTT
TGCTGTATCTGGTGAAACCCACAAAACACCTGAAAAAACTCAAGCTGACTTCCACTGCGAA
GGGAAATTATTGGTTGTGTAACTAGTGGTAGAGTGGCTTCAAGCATAGTTGATCAAAAC
TCCACTCAGTATCTGCATTACTTTATCTGCATAATCTGCATGATAGCTTATTCTCAG
TTATCTTCCCCATAATAAAATATCTGCCAAA

31/246

FIGURE 30

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA125150
><subunit 1 of 1, 505 aa, 1 stop
><MW: 59086, pI: 9.50, NX(S/T): 3
MEPSWLQELMAHPFLLLILLCMSLLLQVIRLYQRRRWIRALHLFPAPPAHWFYGHKEF
YPVKEFEVYHKLMEKYPACAVPLWVGPFMTFFSVHDPDYAKILLKRQDPKSAVSHKILESW
VGRGLVTLDGSKWKKHRQIVKPGFNISILKIFITMMSESVRMMLNKWEEHIAQNSRLELF
QHVSMLTLDSIMKCAFSHQGSIQLDSTLDSYLKAVFNLSKISNQRMNNFLHHNDLVFKFS
SQGQIFSFKFNQELHQFTEKVIQDRKESLKDQLQDTTQKRRWDFLDILLSAKSENTKDFS
EADLQAEVKTFMFAGHDTTSSAISWILYCLAKYPEHQQRCCRDEIRELLGDGSSITWEHLS
QMPYTTMCIKECLRLYAPVVNISRLLDKPITFPDGRSLPAGITVFINIWAHHNPYFWED
PQVFNPLRFSRENSEKIH PYAFIPFSAGLRNCIGQHFAII ECKVAVALTLLRFKLAPDHS
RPPQPVRQVVLKSKNGIHVFAKKVC
```

Important features of the protein:**Signal peptide:**

Amino acids 1-28

Transmembrane domain:

Amino acids 451-470

N-glycosylation sites:

Amino acids 145-149; 217-221; 381-385

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 264-268

N-myristoylation sites:

Amino acids 243-249; 351-357; 448-454; 454-460

Cytochrome P450 cysteine heme-iron ligand signature:

Amino acids 445-455

Cytochrome P450 cysteine heme-iron ligand proteins:

Amino acids 442-473

FAD-dependent glycerol-3-phosphate dehydrogenase proteins:

Amino acids 124-141

32/246

FIGURE 31

TCCGCTGTCGCCAGTCCCAGGCGCTGGCGGGAACTGACCTGGAGCAAGCAGGACCTCCCT
CCCACCTCTCCCGCCTGGCCTCCGCGGGAGTCCCCTACGATCCGCTCAGCAGTGGGCAC
CGCTGAGGACAGCGAGTCCTGGGAGTGAGGCCAAGGCCACCCCTGGCCAGGCCAGGAGAGAT
AGCCAGGGCAGGCCAGCAGCCCAGGCCAGGCTCTGGCCACGGCGGTCTCGACATGGGAGA
GACATTGTCTGCTTTTATCCTGTTAACCTGTCTCGGTGGTGTGCCACGACATTCCCCAG
GGTTCAAGGTGCCCGGTGGCGAGGGTCAGTCCAGTGGTAGAGCCTTGCTCTCCTAGGCTCAT
CCTGCTGGCGGTCTCCTGCTGTGTGGTGTACAGCTGGTTGTGTCCGGTTCTGCT
GCCTCCGGAAGCAGGCCACAGGCCAGGCCACATCTGCCACCAGCACGGCAGCCCTGCGACGTG
GCAGTCATCCCTATGGACAGTGACAGCCCTGTACACAGCACTGTGACCTCCTACAGCTCCGT
GCAGTACCCACTGGGCATGCGGTTGCCCTGCCCTTGGGAGCTGGACCTGGACTCCACGG
CTCCTCCTGCCTACAGCCTGTACACCCCGGAGCCTCCACCCCTACGATGAAGCTGTCAAG
ATGGCCAAGCCCAGAGAGGAAGGACCAGCACTCTCCAGAAACCCAGCCCTCTCCTGGGC
CTCGGGCCTAGAGACCACTCCAGTGCCCTAGGAGTCGGGCCCAATACTCAACTACCACCTT
GTAGCCCTGGTGCCCT**TGA**AGGAGGTAGGAGAACGGACCAGAGCTTGGAGAAACTAATGCTT
GGAGCCAAGGGCCCCAGCCCACCCACCGTCCCACACATTGCTGTGGCCCCAACCTCGGTGC
CATGTTACACCGGCCCTGGCGTCACCCACTAGGCAGGCTGCTGCTTCAGCCTCAGCCCT
GGCCCAAGCCCCAGCAGGCCCTAGCCTGGAAGAGGCCCTGGGCCTAACGCTCGGTGGGA
GCTCAGGGCCACCTGTGACGTCTGCATTTGGAGAGAGAATAAGTTGTATTAAAGTGGT

33/246

FIGURE 32

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA125151
><subunit 1 of 1, 194 aa, 1 stop
><MW: 20882, pI: 6.44, NX(S/T): 0
MERHCLLFILLTCLRWLCHDIPQGSGARPRVSPVVEPCSPRLILLAVLLLLLCGVTAGC
VRFCCCLRKQAAQPHLPPARQPCDVAVIPMDSDSPVHSTVTSYSSVQYPLGMRLPLPFGE
LDLDSTAPPAYSLYTPEPPPSYDEAVKMAKPREEGPALSQKPSPLLGASGLETPVPQES
GPNTQLPPCSPGAP
```

Important features of the protein:

Signal peptide:

Amino acids 1-20

Transmembrane domain:

Amino acids 39-58

N-myristoylation site:

Amino acids 55-61

Prokaryotic membrane lipoprotein lipid attachment site:

Amino acids 50-61

FIGURE 33

CCTTGCTTGGTCTGGCACACACAAATCCAGTGGCTACACAGGTTTCCAGAAGCCCCAC
GAGGTGGTA**ATGGT**GCTGCTGATTCA GACCCCTGGGGCCCTCATGCCCTCGCTGCCCTCTG
CCTCAGCAACGGCGTGGAGAGGGCAGGGCCGAGCAGGAGCTCACCAAGGCTGCTGGAGTTCT
ACGACGCCACCGCCCACCTCGCAAGGGCTTGGAGATGGCACTGCTCCCCACCTACATGAA
ACAATCTGGTAAAAGTCACGGAGCTGGTGGATGCTGTATGATCCATACAAACCTACCAG
CTGAAGTATGGCGACATGGAAGAGAGCAACCTCCATCCAGATGAGTGCTGTGCCTCTGGA
GCATGGGAAAGTGAATGACTGTGCGAGGAGCTGAGCCACTCCGTGAACAAGCTGTTGGTC
TGGCGTCTGCAGCCGTTGACAGATGCGTCAGATTACCAATGGCCTGGGACCTGCGGCCTG
TTGTCAGCCCTGAAATCCCTCTTGCCAAGTATGTTGCTGATTCACCAAGCCTCCAGTC
CATACGAAAGAAGTGCACAACTGGACCACATTCCCTCCAACTCCCTCTCCAGGAAGATTGGA
CGGCTTTCAAGAACTCCATTAGGATAATAGCCACCTGTGGAGAGCTTTGCGGCATTGTGGG
GACTTCGAGCAGCAGCTAGCCAACAGGATTTGTCCACAGCTGGGAAGTATCTATCTGATTC
CTGCAGCCCCCGGAGCCTGGCTGGTTTCAGGAGAGCATCTTGACAGACAAGAAGAACTCTG
CCAAGAACCCATGGCAAGAATATAATTACCTCCAGAAAGATAACCCCTGCTGAATATGCCAGT
TTAATGAAATACTTTACCCCTTAAGGAAAAAGGGTCAAGCAACCACAACCTGCTGGCTGC
ACCTCGAGCAGCGCTGACTCGCTTAACCAGCAGGCCACCAGCTGGCTTCGATTCCGTGT
TCCTCGCAGTCAAACACAGCTGTTGCTTATTCGAAGATGGACAGCTGGAATACGGCTGGC
ATCGGAGAAACCCCTCACAGATGAACTGCCCGCTTACTGCTCACCCCTCGAGTACATCAG
CAACATCGGGCAGTACATCATGTCCTCCCCCTGAATCTTGAGCCATTGTGACTCAGGAGG
ACTCTGCCTTAGAGTTGGCATTGCACGCTGGAAAGCTGCCATTCCCTCTGAGCAGGGGGAT
GAATTGCCCGAGCTGGACAAACATGGCTGACAACGGCTGGCTGGGCTCGATGCCAGAGCCACAAAT
GCAGACCTACTGTGATGCGATCCTACAGATCCCTGAGCTGAGCCACACTCTGCCAAGCAGC
TGGCCACTGACATCGACTATCTGATCAACGTGATGGATGCCCTGGGCTGCAGCCGTCCCGC
ACCCCTCAGCACATCGTACGCTACTGAGAACCCAGGGCTGAGGACTATAGACAGGTAGCAA
AGGCCTGCCCGTCGCCCTGGCACCACCGTGGCACCAGCTGGGAGTGTGAATTACT**TGACCCC**
ACACACACCGGACCACCAAGAGAGCCAGGGCTGCTGTTCTGACTCACCAGCACAGATT
GCTCAGAAACTCTGCCAAGATTGGCAGAAGTTACTTTAAAAAGACTTGGTCAGCTGGTC
ACGGTGGCTACGCCGTAAATCCAGCATTGGGAGGCCAGATGGATCATGAGGCC
AGGAGTTGAGACCAAGCAGCTGACCAACATGGTAAACCCATCTCTACTAAAAATACAAAAAT
TAACAGCAGAGCGAGACTCTGTCAAAAAAAAAAAGACTTGGTCATTGTATAA
TCAAAAAGAGTTGTAATTAAAGATGTATTATTTAGATAATT
TTAAAGGATCAGATCTGAAAATGGAATAACTACTGTGAAATGCAAAA

35/246

FIGURE 34

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA125181
><subunit 1 of 1, 491 aa, 1 stop
><MW: 54759, pI: 5.61, NX(S/T): 0
MVLIIQTLGALMPSLPSCLSNGVERAGPEQELTRILEFYDATAHFAKGLEMAILPHLHEH
NLVKVTELVDavyDPYKPYQLKYGDMEESNLLIQMSAVPLEHGEVIDCVQELSHSVNKL
GLASAAVDRCVRFTNGLGTCGILLSALKSLFAKYVSDFTSTLQSIRKKCKLDHIPNSLFQ
EDWTAFQNSIRIIATCGELLRHCGDFEQQLANRILSTAGKYLSDSCSPRSLAGFQESILT
DKKNSAKNPWQEYNYLQKDNPAEYASLMEILYTLKEKGSSNNLLAAPRAALTRLNQQAH
QLAFDSVFLRIKQQLLISKMDSWNTAGIGETLTDELPAFSLTPLEYISNIGQYIMSLPL
NLEPFVTQEDSALELALAHAGKLPFPPEQGDELPELDNMADNWLSIARATMOTYCDAILQ
IPELSPHSAKQLATDIDYLINVMDALGLQPSRTLQHIVTLLKTRPEDYRQVSKGLPRLA
TTVATMRSVNY
```

Important features of the protein:**Signal peptide:**

Amino acids 1-20

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 242-246

N-myristoylation sites:

Amino acids 22-28; 48-54; 121-127; 136-142; 141-147; 328-334;
447-453

Leucine zipper pattern:

Amino acids 295-317

FIGURE 35

GCAAGTGCCACCATGCTAGTGTATTGGACTTCAGTAAAAGTTAGTTGCTTCCTCCCGT
TGTCCCATCTCACTCCTGGGCCACCCATGGGCTGCTGGTAGCTGGTGTGGCTGCTGCTG
GAETGTGTGGCAGTCATCCATCTGTCAAGCAGCCACTGCAGGCCTACTTGCTGGGTGCCAG
CACCGCACTCACCCTGCAGGGCTGGCCAGGAGCGTAGAGATCCCCAGAGCCATGCCAGTG
AGAGGGCGGCAGGGATAGGTACCCAGGGAATGCCACAGGAGTTGCTGGCTCACGGAGCTC
TTTCACTGGTCAGAGAGGGAGTGTGTAGGAGAGGAGCTTCTACTTGGTGTGAAGGACAGAT
GGGGTTGGCTGGAGAGAGGGAGGAATGTGGCGGGCCTTATAGGCAGGCGAGAAGGTGAGA
GCCAAGGCCCTCTGTGGCAGGGCGAGGTGGCGTGTGAGGAGACTCGTCCAGCTGGCAGA
GGCTCATGT**TGA**GGGATGAGGCAGAGCTGGGGAGGGAGGCCAGAAATGGCAGGTCTT
GAATGCAGGTTGGAAGCAGGGACGCCCTGTGAGGGTACAGAGTCTGGCTGTTACCTCTG
TGGCTTTGCTAGAAGGTGAGATGTCAGGGAGGAAGACAGGACTCCAGGATGTCTCCTGTCT
CTCTGGAAAAAGGAGGTGGCCCTTCTCAGCAGTCAGCTGCTGTTTGAGGTCTTCTCC
ATGGATAATCCACGGTGTGGAAGTGGTAAGGTAATGGATCCTCATGGCTTACCATAAAAA
ATATCTGGAGGCTGGACCATTTCCTTAAACGTTATAAAAGCTGGAATTGAATGCCATCGG
TGTCAACCCCTGGGAAGTGTGCTTCTCTTGAGCTTTGGCCCGAGATAGCAGTCACTCC
ATAGTTCTGTAAGGACCCAGCCTGGTGTGCCTGGTTCTGCCATTAGGGAGCAGCTAGAGG
TCTCCAGTAGCTCCTGTAAAGTGTGAAAGAAAAGGGCTGGGTGCTGACTGCTCCTGGA
GAAAAGCAACACACTCCAAAGTCTTAATTGCCTGCTCCAGGGAGCTGTGGTGGTTCCCT
TGGCAGGGCACACGCCAGTGGTTGACTTAATAAGGATACATTAAATCAGAGGACAAAAA
ATGTGCCCTGACTTGATTCCGCATGGCTCCAGCATGGTCAAAGG

37/246

FIGURE 36

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA125192
><subunit 1 of 1, 139 aa, 1 stop
><MW: 14841, pI: 9.20, NX(S/T): 0
MGLLVAGVWLLDCCAVHPSVSSHCGPTCWVPSTALTTAGVARSVRSPEPMASERRPGIG
TQGMPQEFAGLTELFWSERSVCRRGLLGVEGQMGFGWERGGMWAGLIGRREGESQGPL
WAGRGGVLRRLVQLRGSC
```

Important features of the protein:

Signal peptide:

Amino acids 1-22

N-myristoylation sites:

Amino acids 2-8;40-46;86-92;102-108;103-109

Amidation site:

Amino acids 109-113

38/246

FIGURE 37

GGCCAGGA**ATGGGGTCCCCGGGCATGGTCTGGGCCTCCTGGTGCAGATCTGGCCCTGCAA**
GAAGCCTCAAGCCTGAGCGTGCAGCAGGGGCCAACTTGCTGCAGGTGAGGCAGGGCAGTCA
GGCAGCCCTGGTCTGCCAGGTGACCAGGCCACAGCCTGGAACGGCTCCGTGTTAAGTGGACA
AAGGATGGGCCATCCTGTGTCACCGTACATCACCAACGGCAGCCTCAGCCTGGGGTCTG
CGGGCCCCAGGGACGGCTCTCTGGCAGGCACCCAGCCATCTCACCTGCAGCTGGACCCCTG
TGAGCCTCAACCACAGCGGGCGTACGTGTGCTGGCGGCCGTAGAGATTCTGAGTTGGAG
GAGGCTGAGGGCAACATAACAAGGCTTTGTGGACCCAGATGACCCCACACAGAACAGAAA
CCGGATCGCAAGCTCCAGGATTCCCTTCGTGCTGCTGGGGTGGGAAGCATGGGTGTTG
CTGCGATCGTGTGGGTGCCTGGTCTGGGCCGCGCAGCTGCCAGCAAAGGACTCAGGA
AATGCATTCTACAGCAACGTCTATACCGGCCCCGGGGCCCAAAGAAGAGTGAGGACTG
CTCTGGAGAGGGGAAGGACCAGAGGGGCCAGAGCATTATTCAACCTCTCCGCAACCGG
CCCCCGCCAGCCGACCTGGCGTCAAGACCTGCCAGCCGAGACCCCTGCCAGCCCC
AGGCCCGGCCACCCGTCTCTATGGTCAGGGTCTCTCTAGACCAAGCCCCACCCAGCAGCC
GAGGCCAAAAGGGTCCCCAAAGTGGAGAGGAGT**GA**AGAGATCCCAGGAGACCTCAACAGGA
CCCCACCCATAGGTACACACAAAAAGGGGGATCGAGGCCAGACACGGTGGCTCACGCCCTG
TAATCCCAGCAGTTGGGAAGCCGAGGCCGGTGGAACACTTGAGGTGAGGGTTGAGACCA
GCCTGGCTTGAACCTGGAGGCCGGAGGTGCAAGTGAGCCGAGATTGCGCCACTGCACTCCAG
CCTGGCGACAGAGTGAGACTCCGTCTCAAAAAAAACAAAAGCAGGAGGATTGGAGCC
TGTCAAGCCCCATCTGAGACCCGTCCTCATTTCTGTAATGATGGATCTCGCTCCACTTTC
CCCCAAGAACCTAATAAGGTTGTGAAGAAAAAAAAAAAAAA

FIGURE 38

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA125196
><subunit 1 of 1, 278 aa, 1 stop
><MW: 30319, pI: 9.21, NX(S/T): 3
MGSPGMVLGLLVQIWAHQEASSLSVQQGPNLLQVRQGSQATLVCQVDQATAWERLRVKWT
KDGAILCQPYITNGSISLGVCGPQGRLSWQAPSHLTQLQDPVSLNHSGAYVCWAIVEIPE
LEEAEGNITRLFVDPDDPTQNRNRRIASFPGFLFVLLGVGSMGVAAIVWGAWFWGRRSCQQ
RDSGNAFYSNVLYRPRGAPKKSEDCSGEGKDQRGQSIYSTSFPQPAPRQPHLASRPCPSP
RPCPSPRPGHPVSMVRVSPRSPPTQQPRPKGFPKGEE
```

Important features of the protein:**Signal peptide:**

Amino acids 1-22

Transmembrane domain:

Amino acids 149-166

N-glycosylation sites:

Amino acids 73-77;105-109;127-131

Glycosaminoglycan attachment site:

Amino acids 206-210

N-myristoylation sites:

Amino acids 5-11;37-43;63-69;108-114

Amidation site:

Amino acids 173-179

FIGURE 39

ACCAAGCAGAAGGCTGGGAGTCTGTAGTTGTTCCCTGCTGCCAGGCTCCACTGAGGGGAACGG
GGACCTGTCTGAAGAGAAGATGCCCTGCTGACACTCTACCTGCTCCTCTGGCTCTCAG
GCTACTCCATTGCCACTCAAATCACCGGTCCAACAAACAGTGAATGGCTGGAGCAGGGGCTCC
TTGACCGTGCGAGTGTGTTACAGATCAGGCTGGGAGACCTACTTGAAGTGGTGGTGTGAGG
AGCTATTTGGCGTGACTGCAAGATCCTGTTAAAACCAGTGGGTAGAGCAGGAGGTGAAGA
GGGACCGGGTGTCCATCAAGGACAATCAGAAAAACCGCACGTTCACTGTGACCATGGAGGAT
CTCATGAAAAGTGTGACACTTACTGGTGTGAAATTGAGAAAATGGAAATGACCTTGG
GGTCACAGTCAAGTGACCATTGACCCAGCACCAGTCACCCAAAGAAGAAAATAGCAGCTCCC
CAACTCTGACCGGCCACCACTTGGACAACAGGCACAAGCTCCTGAAGCTCAGTGTCCCTCG
CCCCTCATCTTACCATATTGCTGCTGCTTTGGTGGCCGCTCACTCTGGCTGGAGGATG
ATGAAGTACCAGCAGAAAGCAGCCGGATGTCCCCAGAGCAGGTACTGCAGCCCTGGAGGG
CGACCTCTGCTATGCAGACCTGACCCCTGCAGCTGGCCGAACTCCCCCGCAAAGGCTACCA
CGAAGCTTCCCTCTGCCAGGTGACCAGGTGGAAGTGGAAATATGTCACCATGGCTTCCCTG
CCGAAGGAGGACATTTCCTATGCATCTGACCTGGGTGCTGAGGATCAGGAACCGACCTA
CTGCAACATGGGCCACCTCAGTAGCCACCTCCCCGGCAGGGCCCTGAGGAGCCCACGGAAT
ACAGCACCATCAGCAGGCCTTAGCCTGCACCTCAGGCTCCTCTGGACCCCAGGCTGTGAG
CACACTCCTGCCTCATCGACCGTCTGCCCTGCTCCCTCATCAGGACCAACCCGGGACT
GGTGCCTCTGCCTGATCAGCCAGCATTGCCCTAGCTCTGGTTGGCTGGGCCAAGTCT
CAGGGGGCTCTAGGAGTTGGGTTTCTAAACGTCCCCTCTCCCTACATAGTTGAGGAG
GGGGCTAGGGATATGCTCTGGGCTTCATGGGAATGATGAAGATGATAATGAGAAAAATGT
TATCATTATTATCATGAAGTACCAATTATCATAATACAATGAACCTTATTATTCATTGCCTACCA
CATGTTATGGGCTGAATAATGGCCCCAAAGATACTGTGCTCTAATCCTCAGAACTGTGA
CTGTTACCTCTGTGGCAGAAAGGGACAGTGCAGATGTATGTAAGTTAAGGACTTTGAGATA
GAGAGGTTATTCTTGTGATTCAAGTGGGCCAAATATCACCACAAGGGCTCTCATAGAA
AGAGGCCAGAAGGTCAAAGAGGTAGAGACAAAGTGTGATGGAAGTGGACGTGGGTGTGACG
TGAGCAGGGCCATGAATGCCAGCCTCAGATGCCAGAAAGGGAAAGGAATGGATTCCCC
TGCCTGGAGGCTCCAAAAGAAACCAAGCCCTGCCACGCCCTGACTTGAGGCCATTGAAACTG
ATCTTGAGCTCCTGGCCTCCAGAATTGCAGGAGAATAATTGTGTTTTAATGAAAAA
AAG

FIGURE 40

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA125200
><subunit 1 of 1, 290 aa, 1 stop
><MW: 32335, pI: 5.82, NX(S/T): 1
MPLLTLYLLLFWLSGYSIATQITGPTTVNGLERGSLTVQCVYRSGWETYLKWWCRGAIWR
DCKILVKTSQEVKDRVSIKDNQKNRTFTVTMEDLMKTDADTYWCGIEKTGNDLGVT
VQVTIDPAPVTQEETSSSPTLTGHHLDRHKLLKLSVLLPLIFTILLLLVAASLLAWRM
MKYQQKAAGMSPEQVLQPLEGDLCYADLTLQLAGTSRKATTKLSSAQVDQVEVEYVTMA
SLPKEDISYASLTGAEQEPTYCNMGHLSSHLPGRGPEEPTEYSTISRP
```

Important features of the protein:**Signal peptide:**

Amino acids 1-15

Transmembrane domain:

Amino acids 155-174

N-glycosylation site:

Amino acids 88-92

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 218-222

Tyrosine kinase phosphorylation site:

Amino acids 276-285

N-myristoylation sites:

Amino acids 30-36;109-115;114-120

FIGURE 41

AAGAACACTGTTGCTTGGACGGGCCAGAGGAATTAGAGTAAACCTGAGTCCT
GGTCCGTGAGAATTCAAGC**ATG**GAATGTCCTACTATTCTGGATTCCTGCTGGCTG
CAAGATTGCCACTTGATGCCCAAACGATTCACTGATGTCCTGGCAATGAAAGACCTCT
GCTTACATGAGGGAGCACAATCAATTAAATGGCTGGTCTCTGATGAAAATGACTGGAATGA
AAAATCTACCCAGTGTGGAAGCAGGGAGACATGAGGTGGAAAAACTCTGGAAGGGAGGCC
GTGTGAGGCGGTCTGACCAAGTGAATCACCAGCCTGTCGGCTCAAATATAACATTGCG
GTGAACCTGATATTCCCTAGATGCCAAAGGAAGATGCCATGGCAACATAGTCTATGAGAA
GAATGCAAGAAATGAGGCTGGTTATCTGCTGATCCGTATGTTACAACGGACAGCATGGT
CAGAGGACAGTGACGGGAAATGGCACCGCCAAAGCCATCATAACGTCTCCCTGATGGG
AAACCTTCTCACCACCCGGATGGAGAAGATGGAATTCTACATCTACGTCTCACACACTT
GGTCAGTATTCCAGAAATTGGGACGATGTCAGTGAGAGTTCTGTAACACAGCCAAATGT
GACACTTGGGCCTCAACTCATGGAAGTGACTGTCACAGAAAGACATGGACGGCATATGTC
CCATCGCACAAGTGAAGAGATGTCAGTGGTAACAGATCAGATTCTGTGTTGTGACTATG
TTCCAGAAGAACGATGAAATTCACTCGACGAAACCTTCTCAAAGATCTCCCCATTATGTT
TGATGTCCTGATTCACTGATCTAGCCACTTCTCAAATTATTCTACCATTAACTACAAGTGG
GCTTCGGGATAATACTGGCCTGTTGTTCCACCAATCATACTGTGAATCACACGTATGT
CTCAATGGAACCTTCAGCTTAACCTCAACTGTGAAAGCTGCAGCACCAGGACCTGTCCGCC
ACCGCCACCACCACCCAGACCTCAAAACCCACCCCTTTAGCAACTACTCTAAATCTT
ATGATTCAAACACCCAGGACCTACTGGTGAACACCCCTGGAGCTGAGTAGGATTCTGAT
GAAAATGCCAGATTAACAGATATGGCACTTCAAGCCACCATCACAATTGTAGAGGGAAAT
CTAGAGGTTAACATCATCCAGATGACAGACGTCCTGATGCCGGTGCCATGGCTGAAAGCT
CCCTAATAGACTTGTGCGTGAACCTGCCAAGGGAGCATTCCCACGGAGGTCTGTACCATCATT
TCTGACCCCACCTGCGAGATCACCAGAACACAGTCTGCAGCCCTGTTGAGTGTGAGAT
GTGCTGACTGTGAGACGAACCTCAATGGGCTGGGACGTACTGTGAAACCTCACCC
TGGGGATGACACAAGCTGGCTCAGAGCACCCCTGATTTCTGTTCTGACAGAGACCA
GCCTCGCCTTAAGGATGGCAAACAGTGCCTGATCTCCGTTGGCTGTTGGCATATTGT
CACTGTGATCTCCCTTGGTGTACAAAAACACAAGGAATACAACCAATAGAAAATAGTC
CTGGGAATGTGGTCAGAACAGCTGAGTGTCTTCTCAACCGTGCACAGCCGTGTC
TTCCCGGAAACCAGGAAAGGATCCGCTACTCAAAACCAAGAATTAAAGGAGTTCT**TA**
ATTTGACCTGTTCTGAAGCTCATTTCAGTGCATTGATGTGAGATGTGCTGGAGTG
GCTATTAAACCTTTCTCAAAGATTATTGTTAAATAGATATTGTGTTGGGAAAGTTGA
ATTTTTATAGGTTAAATGTCATTAGAGATGGGAGAGGGATTATACTGCAGGCAGCTC
AGCCATGTTGAAACTGATAAAAGCAACTTAGCAAGGCTCTTTCTATTATTTTATGTT
TCACTTATAAAAGTCTAGGTAACTAGTAGGATAGAAACACTGTGTCCTGAGAGTAAGGAGAG
AAGCTACTATTGATTAGAGCCTAACCCAGGTTAAGTCAAGAACAGGGGGATACTTCA
TTCCATGTAAGTGCATAAGCCAATGTAGTCCAGTTCTAAGATCATGTTCCAAGCTA
ACTGAATCCCACCTCAATACACACTCATGAACTCCTGATGGAACAATAACAGGCCAAGCCT
GTGGTATGATGTGCACACTGCTAGACTCAGAAAAAAACTACTCTCATAAATGGGTGGAG
TATTTGGTACAACCTACTTGTGCTGGCTGAGTGAAGGAATGATATTCAATATTCA
TTCCATGGACATTAGTTAGTGTCTTTATATACCAAGGCATGATGCTGAGTGACACTCTGT
GTATATTCCAATTTGTACAGTCGCTGCACATATTGAAATCATATATTAAAGACTTCC
AAAGATGAGGTCCCTGGTTTCACTGGCAACTTGATCAGTAAGGATTCAACCTCTGTTGA
ACTAAAACCATCTACTATATGTTAGACATGACATTCTCTCCTGAAAGAGACA
GTGTGGAAAGAGACA

FIGURE 42

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA125214
><subunit 1 of 1, 572 aa, 1 stop
><MW: 63953, pI: 6.55, NX(S/T): 12
MECLYYFLGFLLLAARLPLDAAKRFHDVLGNERPSAYMREHNQLNGWSSDENDWNEKLYP
VWKRGDMRKNSWKGRVQAVLTSALVGSNITFAVNLI FPRCQKEDANGNIVYEKNC
RNEAGLSADPYVYNWTAWSEDSDENGTGQSHHNVPDGKPFPHPGWRRWNFIYVFHTL
GQYFQKLGRCVRVSVNTANVTLGPQLMEVTYRRHGRAYVPIAQVKDVTYVVTDQIPVVF
TMFQKNDRNSSDETFLKDLPIMFVDVLIHDPSHFLNYSTINYKWSFGDNTGLFVSTNHTVN
HTYVLNGTFSLNLTVKAAAPGPCPPPPPPRPSKPTPSLATTLKSYDSNTPGPTGDNPLE
LSRIPDENQINRYGHFQATITIVEGILEVNI IQMTDVLMPPWPESSLIDFVVTQGSIP
PTEVCTIISDPTCEITQNTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTGDDTSALTS
TLISVPDRDPASPLRMANSALISVGCLAIFVTVISLLVYKKHKEYNPIENSPGNVVRSKG
LSVFLNRAKAVFFPGNQEKDPLKNQEFKGVS
```

Important features of the protein:**Signal peptide:**

Amino acids 1-21

Transmembrane domain:

Amino acids 496-516

N-glycosylation sites:

Amino acids 93-97;134-138;146-150;200-204;249-253;275-279;
296-300;300-304;306-310;312-316;459-463;467-471

N-myristylation sites:

Amino acids 91-97;147-153;290-296;418-424

Prokaryotic membrane lipoprotein lipid attachment site:

Amino acids 496-507

Cell attachment sequence:

Amino acids 64-67

FIGURE 43

45/246

FIGURE 44

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA125219
><subunit 1 of 1, 283 aa, 1 stop
><MW: 31175, pI: 7.51, NX(S/T): 0
MADPHQLFDDTSSAQSRGYGAQRAPGGLSYPAASPTPHAAFLADPVSNMAMAYGSSLAAQ
GKELVDKNIDRFIPITKLKYYFAVDTMYVGRKLGILFFPYLHQDWEVQYQQDTPVAPRFD
VNAPDLYIPAMAFITYVLVAGLALGTQDRFSPDLLGLQASSALAWLTLEVLAILLSLYLV
TVNTDLTTIDLVAFLGYKYVGMIGGVLMGLLFGKIGYYLVLGWCCVAIFVFMIRTLRLKI
LADAAAEGVPVRGARNQLRMYLTMAVAAQPMLMYWLTFLV
```

Important features of the protein:**Transmembrane domain:**

Amino acids 126-142;164-179;215-233

N-myristoylation sites:

Amino acids 54-60;141-147;156-162;201-207;205-211;209-215

Amidation site:

Amino acids 89-93

46/246

FIGURE 45

GCTGAGCACCAACAGGAACATTCCAGTGAAGAGCAAGTGCTGCCGACCCAGGACCCTGTG
CCAGGCTGGCAGCCCTCCAGCTCCCTCAGAGAGGAAACCTCTGTCTGGCTGAGGGTGGGAC
TAGCTGGG**ATGT**CTCACTCCAGTTGCTCAGGTCACCCAGGAAGCTCCTCCGTGGAGTGGCC
AGCCTGATTCTAGCCCTGTCCCTCTGGCAGCACATGCCACACCTGCCTGGGCCTCTGCTC
CCTGATGCTGTGATGAGCCCCTGCCCTCAATGTTCTCAAAGACAGACCCCCCTGAGGCCAGC
TTGAATGTGAAGACTGCTGAAGTCAGCTGGCTTCACTTGAGCTGCAGAAAAGGTGGCTGGGA
TGGCCCAGGTGCACCCAGAGGCCCCAGGCCCTTGCTGCCTTGGGT**TGAC**TTGGTTGT
CTCTGAGGCCCTGCCAGAGCTGGGCTGCCGGTGGTGGCGGTCCGACCTCGGGCAGTCAGT
GCTCCGCAGCCTCAGCACTGCATCCCAGACCCAGTGTCCCTCAGAGGGAAAGGCCAGCCTCCC
TGCCTCATGGAACCAGGAGTCCAAAAAGTCAGGAGCCTGGAGGCTCTGAAAGGAGCAGGGA
TTCCATAGTGCCTGAAGCTGAAATAGGCCTGCCTGGGGAGCCCCCAGCAAAACTGTTTT
CATACCCACTCCCAGAACTGCCCGCTCCAGCTCCAGCGCCAGCGCCAGCTGGTTGCCAGGC
GTCATTGGAGAGGCCCTGGCTGCCAGGGGAGCAGGGAGTGGTGGACCTGTATGGGCTGGC
AGGAGGCCATTGGCATGCTGACAAGTGTACCTGCCTCCAGCTGGGAGCAGGCCACCCCTCAG
GTGGCCTGCTTGCACCTCCTATCCGGAGGTAGCCTGCCACCTGTAGGCAGAGGGGGCTCT
TGCTTGAGGCCTGCACAGGAAGCAAGTATAGCCCCGGTGCCCCAGAGTGGTTCCACTAGC
CCTGGCGAGATGGCCTGCTGAGATCTGCTCCAGACCCCACCATCTGGGAGCACAGT
CCTTAGGCTGCCTGGTCCAGGAAGGGGGTGCCTGTCAAGGAAACCTGGACTCTCAAGGC
CCACCAGCCTCTCCGTGAGTGTAGAAATCACAGATACTATACCTAATTACACTACTC
ACTACTCAAAAAAAAAAAAAAA

47/246

FIGURE 46

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA128309
><subunit 1 of 1, 97 aa, 1 stop
><MW: 10112, pI: 8.64, NX(S/T): 0
MSHSSCSGSPRKLLRGVASLILALSSLAHATPAWAFCSMLDEPLPPQCFSKTDPEAS
LNVKTAEVSWLHLSCRKGWDGPGAPRGPSPLAAFGL
```

Important features of the protein:

Signal peptide:

Amino acids 1-31

48/246

FIGURE 47

TTCCGGGCCCTGGCGTCTCGTCTCCTTACCCCTGGGGCTACCCCTGCCCGGTCTACTGCCCG
CGGTTAACCGCCGCGAGCCGCCCTCCCCCTCCCCGCCGACTCAACCCCTGCCCTCCCCCGT
GCTTGAGACGCCGCCGGGGGCCAGGGGGCTGATATGCGTGGGCCTCGCGCTGATCTTG
GTGGGCCACGTGAACCTGCTGCTGGGGCGTGCTGCATGGCACCGTCCCTGCGGCACGTGGC
CAATCCCCGCGCGCTGTACGCCGGAGTACACCGTAGCCAATGTCATCTGTGCGGCTCGG
GGCTGCTGAGCGTTCCTGGGACTTGTGGCCCTCCTGGCGTCCAGGAACCTCTTCGCCCC
CCACTGCACTGGGTCCCTGCTGGCACTAGCTCTGGTGAACCTGCTCTTGTCCGTTGCCTGCTC
CCTGGGCCTCCTCTGCTGTCACTCACTGTGCCAACGGTGGCCGCCCTATTGCTG
ACTGCCACCCAGGACTGCTGGATCCTCTGGTACCACTGGATGAGGGGCCGGGACATACTGAC
TGCCCCCTTGACCCCACAAGAATCTATGATACAGCCTTGGCTCTGGATCCCTCTTGCT
CATGTCTGCAGGGGAGGGCTGCTCTATCTGGTTACTGCTGTGGCTGCACACTACAGTG
GAGTTGGGCCCTGCAGGAAGGACGGACTTCAGGGCAGCTAGAGGAATGACAGAGCTTGAA
TCTCCTAAATGTAAGGCAAGGAAATGAGCAGCTACTGGATCAAAATCAAGAAATCCGGGC
ATCACAGAGAAGTTGGGTTTAGGACAGGTGCTGTCCGAGACTCAGTCCCTAAAGGGTTTT
TTCCCACTAAGCAAGGGGCCCTGACCTCGGGATGAGATAACAAATTGTAATAAGTAACCTC
TCTTTCTTCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 48

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA129535
><subunit 1 of 1, 222 aa, 1 stop
><MW: 23566, pI: 6.70, NX(S/T): 0
MRVGLALILVGHVNLILGAVLHGTVLRHVANPRGAVTPEYTVANVISVGSGLLSVSGLV
ALLASRNLLRPLHWVLLALALVNLLSVACSLGLLLAVSILTANGGRRLIADCHPGLLD
PLVPLDEGPGHTDCPFDPTRIYDTALALWIPSLLMSAGEAALSGYCCVAALTLRGVGPCR
KDGLQGQLEEMTELESPKCKRQENEQLLDQNQEIRASQRSWV
```

Important features of the protein:

Signal peptide:

Amino acids 1-18

Transmembrane domain:

Amino acids 44-60;76-96

N-myristoylation sites:

Amino acids 94-100;175-181

Amidation site:

Amino acids 106-110

Prokaryotic membrane lipoprotein lipid attachment site:

Amino acids 81-92

50/246

FIGURE 49

CGTCAGTCTAGAAGGATAAGAGAAAGAAAGTTAACGCAACTACAGGAAATGGCTTGGGAG
TTCCAATATCAGTCTATCTTATTCAACGCAATGACAGCACTGACCGAAGAGGCAGCCG
TGACTGTAACACCTCCAATCACAGCCCAGCAAGCTGACAACATAGAAGGACCCATAGCCT
TGAAGTTCTCACACCTTGCCTGGAAGATCATAACAGTTACTGCATCAACGGTGCTTGTG
CATTCCACCATGAGCTAGAGAAAGCCATCTGCAGGTGTTACTGGTTACTGGAGAAA
GGTGTGAGCACTTGACTTTAACCTCATATGCTGTGGATTCTTATGAAAAAATACATTGCAA
TTGGGATTGGTGTGATTACTATTAAGTGGTTCTTGTATTTTTACTGCTATATAA
GAAAGAGGTATGAAAAAGACAAAATATGAAGTCACCTCATATGCAATCGTTGACAATA
GTATTCAAGGCCCTATAATGTGTCAGGCACTGACATGTAAAATTTTTAATTAAAAAAG
AGCTGTAATCTGGCAAAAAGTTCTATGTAATATTTTCAATGCCTTCTCATAAACCCA
GACGAGTGGTAAAATTGCCTTCAGTTGTAATAGGAGAGTTCAAACGTACAGTCTCCCT
TCAACCTATCTGTCTGCCCATATCAAATTATAAATGAGGAGGACAGCAGGCCAAG
AAAGTAGGGACTAAGTATGTCTTGTCAAAATTGTATATTCACTGACTTACACTATGCCT
AGCACACACACACTGAGTAAATATTGTTGAGTGAAATAAAATCAAGAAACAAGTAA
AAACTGA

FIGURE 50

>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA129549
><subunit 1 of 1, 133 aa, 1 stop
><MW: 14792, pI: 5.97, NX(S/T): 0
MALGVPISVYLLFNAMTALTEAAVTVTPPITAQQADNIEGPIALKFSHLCLEDHNSYCI
NGACAFHHELEKAICRCFTGYTGERCEHLLTSYAVDSYEKYIAIGIGIVGLLLSGFLVIF
YCYIRKRYEKDKI

Important features of the protein:**Signal peptide:**

1-20 (weak)

Transmembrane domain:

103-117

N-myristoylation site.

4-10;106-112;110-116

EGF-like domain cysteine pattern signature.

75-87

Integrins beta chain cysteine-rich domain proteins

66-88

FIGURE 51

GGCTCGAGCTGGCTCTCAGACCATCCTGGTGGAAAGAAACACTAGCAGTCTGCCAATCTGA
ATGCAAATCCAGAATAATCTTTCTTTGTTACACAGTATGAGTGCAATTAAATG
GCTGCTACTCTACAGCCTGCCTGCCTTATGCTTCTCCTGGCACGCAGGAAAGTGAGAGCT
TCCACTCCAAAGCAGAGATCCTAGTGACACTAAGTCAGGTATAATCTCCAGCTGGACCT
CATGCACTCACATGGACAACACACTCTCCTCAGTGATCATCATCCTGTACCATGTTG
GTGGCATGCTGTAATCGTGAUTCAACATCCGGTTGCCATTGCTATGTAACAAACCACTCA
ACATTCACTGGCTTGAATTGAAAGCAGGGCTTGAAGAGATATTGCACATTCATCCTCCC
AGCAGCATTATTCAACACAGCCAATAGGCAGAAGCAACCCAAATGTCCAACCATAGATGAGTG
GATAACCAAAATGTAGTCCATCCATACAATGAAATATGATTCAACAGGAAGGAAG
TCCCGCCACGTGCTACAACATGGATGGACCTTGAGGACACTATGCTAAGTGAAGTAAGCCAG
GCACAAAAGGACAATACTCTATGATTCCATTATAGGGTACCAAAGAGAATCAAACCTCAC
AGAGATAGAAAGTAGACTGGGTGGCCAGGGACTCGGGGAGAGAGGAAAGGGCAGTTATTGT
TTAAAAGGTACAGAGTTCAAGTTGGGAAGATGAAATGTTCTGGAAACGGTTATGGTGT
TTTACATTGTTATGTTACGATTTGAAAGAGCAGCTGCCTGAGATTTCACTGGCTTAAAGAAGACGGCAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

53/246

FIGURE 52

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA129580
><subunit 1 of 1, 114 aa, 1 stop
><MW: 12886, pI: 7.04, NX(S/T): 0
MQIQNNLFFCCYTVMSAIFKWLLYSLPALCFLLGQTQESESFHSKAEILVTLSQVIISPA
GPHALTWTTHFSPSVIIILVPCWWHAVIVTQHPVANCYVTNHLNIQWLELKAGS
```

Important features of the protein:

Signal peptide:

Amino acids 1-33

Transmembrane domain:

Amino acids 71-86

N-myristoylation site:

Amino acids 35-41

FIGURE 53

TTTGAAATGGTT **ATG**ACCTCTCCCCACTCCCCGCTTGCTCATTAGTGTCCCTA
GGTGGCTGCTGGGTGACGGGCTTTCATCATCTGATGTGGGCCAGTCGAAAGAGCAGCT
GCAACATCTGTTCTAATTGGGCGCCTTATAAAACTCTTGCTCTTGTACATTGTCACATTG
CTTCCCTCCCACCCGTCTTCCTGGAGACTGCAGAATCTGTAAGCGTCCCTGGAATGCAC
ACGTGGACCTTGTCACTCCAAACAGACTTCTGCTGGTCAGCACTTGTAAATGTCGGCTG
TTACAGGCATT**AGT**CACTTGTGCTCAGAGAGAGACTGTGGTCTTGGAAACTGAAGAAAATGTC
TTTTTGTGTTGTTAATTCTTGGCATCCAGTTAGATTTAACTTCTCAAGAGTTACACAGA
CTTTAGAAAAACATTCTGTCTAAGAAAAAGTGTCTAGCTTGTACAGTTGGATT
TTCACACTTGGTGGTGTGTTGCTGAAATGCTGTTGCTAGTGAATTCCCCCTCCCCCTAT
CTGGGTTGTAAGCAGCTCTGGGCTCTGTCACCTCGGATACCTGTTCTGGGACTGCTT
TTCAACAGCGTTTCTAAGGGCATATGAGAAATTAAATTCTGATGGAATGAAGGTGAAA
CTCTAGTCCCAGGTAAACCTGGTAGGCTGTAGAGACAGAAAGGGGCTGCAGGTCTAGGTG
GAAGAACGAGAACGAATGCAGCATGGTATTCCAGGCCCTTGTAGATTCGGCTTCATCCACAA
CCAATGTGAGTTCTTATCTGCAAAGCGGGCTAAGTGTAAATGGAGGGAAAGGTGGCCAGGCA
CCAGGGTCTGGTTCTCCGCGCCTCACTCTGTCCTCACCTGGCCATGCATAAAAGAACAC
TAGTCAAGTAGCCATTGTACCTGTTCTTATCTGAAAATGAGAAGGTGGAGAGTATGACT
TCTGTTGAAACAACAAGGCCTTACAAATTGGTGAAGTCGAATGAGGGCAGCGTTAAGAG
AAATATCAAAGTTAGTCATTGGATTTCAAGGGCTTAGGGATGGAAACCAGCTGGTAGACT
GGTTGTAGTTATGTCCAAGGGCAGAGTGGAAAAATTGGCCAAAAGAGTGTGGTGGGTG
ACCAGCAAATGTTAGAGGTATACATCAGGGCACAGAGGAGAAAAGCTAACATGATACTGATG
ACTTCAAGTCTTCACTGTCATTCAAGAGGATAGGGGAGGGTTAACGCTGATTAAACAGTGG
GCTTTTTCTCCTGCAAGAGGGTGGAGGTCTATAACTGTGCAGATTATCAGATGCATGC
TAATACATGTTATTCTGGGGACTCTCTTACCTGAAAGTAGACATTGCTGCTATTGCGT
GAAAAAAATAGGAGGACTTATTGAGAATGGGATAGGCTGAGTCCACCGAGATGT
TGGCTTAGAGATGCCGGCATGCTGTACAGTAGGAAGCCCAGCAGAGGAGATTGGCTGT
GTGGGTGATGACAAAGGGAGTTGTTAGCTTATGGTTGGCTATTAAAGTCATGGCAAGGATG
GGCAAGAAAAGTGTAAAATGAGCTGACAAAAGATAATGTTAATTA

55/246

FIGURE 54

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA129794
><subunit 1 of 1, 102 aa, 1 stop
><MW: 11382, pI: 8.72, NX(S/T): 0
MTSSPLPRLLCQLVFLGGCWVTGFSSSLWASAKEQLQHLFLIGSCLYFLPICHIASL
PPCLPWSTAESVSPGMHTWTLSFPNRLSAGQHFVMFGCYRH
```

Important features of the protein:**Signal peptide:**

Amino acids 1-21

N-myristoylation site:

Amino acids 18-24

Prokaryotic membrane lipoprotein lipid attachment sites:Amino acids 9-20; 36-47;
89-100

FIGURE 55

ACACTGGCAAACACTCGCATCCCAGGGCGTCTCCGGCTGCTCCCATTGAGCTGTCTGCTCG
CTGTGCCGCTGTGCCTGCTGTGCCCGCGCTGTCGCCGCTGCTACCGCGTCTGCTGGACGCG
GGAGACGCCAGCGAGCTGGTGATTGGAGCCCTGCGGAGAGCTCAAGCGCCAGCTCTGCCCG
AGGAGCCCAGGCTGCCCGTGAGTCCCAGTTGCTGCAGGAGTGGAGCC**ATGAGCTGCGTC**
CTGGGTGGTGTGTCATCCCCTGGGGCTGCTGTTCTGGTCTGCGGATCCAAGGCTACCTCCT
GCCCAACGTCACTCTTAGAGGAGCTGCTCAGCAAATACCAAGCACACGAGTCTCACTCCC
GGTCCCGAGGCCATCCCAGGGAGGACAAGGAGGAGATCCTCATGCTGCACAACAAGCTT
CGGGGCCAGGTGCAAGCCTCAGGCCCTAACATGGAGTACATGACCTGGGATGACGAACCTGGA
GAAGTCTGCTGCAGCGTGGGCCAGTCAGTGCATCTGGGAGCACGGGCCACCAGTCTGCTGG
TGTCCATGGGCAGAACCTGGCGCTCACTGGGGCAGGTATCGCTCTCCGGGGTTCCATGTG
CAGTCCTGGTATGACGAGGTGAAGGACTACACCTACCCCTACCGAGCGAGTGCACCCCTG
GTGTCCAGAGAGGTGCTCGGGCCTATGTGCACGCACTACACACAGATAGTTGGGCCACCA
CCAACAAAGATCGGTTGTGCTGTGAAACACCTGCCGAAAGATGACTGTCTGGGAGAAGTTGG
GAGAACCGGGTCTACTTGTCTGCAATTATTCTCAAAGGGGAACTGGATTGGAGAACCCCC
CTACAAGAAATGGCCGCCCTGCTCTGAGTGCCAACCCAGCTATGGAGGCAGCTGCAGGAACA
ACTTGTGTTACCGAGAAGAAACCTACACTCCAAACCTGAAACGGACGAGATGAATGAGGTG
GAAACGGCTCCCATTCTGAAGAAAACATGTTGGCTCCAACCGAGGGTATGAGACCCAC
CAAGCCAAGAAAACCTCTGGGTCAACTACATGACCCAAAGTCGTCAAGATGTGACACCAAGA
TGAAGGACAGGTGCAAAGGGTCCACGTGTAACAGGTACCAAGTGCAGGCTGCCTGAAC
CACAAGGCGAAGATCTTGAAGTCTGTCTATGAAAGCTCGTCTAGCATATGCCGCGCCGC
CATCCACTACGGGATCCTGGATGACAAGGGAGGCCTGGTGGATATCACCAGGAACGGGAAGG
TCCCCTCTCGTGAAGTCTGAGAGACACGGCGTGCAGTCCTCAGCAAATACAAACCTTCC
AGCTCATTGTTGGTGTCAAAGTGAAGTGCAGGATTGGACTGCTACACGACCGTTGCTCA
GCTGTGCCCGTTGAAAAGCCAGCAACTCACTGCCAAGAATCCATTGTCGGCACACTGCA
AAGACGAACCTTCTACTGGCTCCGGTTGGAACCAACATCTATGCAAGATACCTCAAGC
ATCTGCAAGACAGCTGTGCACGGGGAGTCATCAGCAACGAGAGTGGGGGTGACGTGGACGT
GATGCCCGTGGATAAAAAGAAGACCTACGTGGCTCGTCAGGAATGGAGTTCAAGTCTGAAA
GCCTGGGACTCCTCGGATGAAAGGCCCTCGGATCTTGCTGTCAGGCAGT**GA**ATTCC
AGCACCAGGGGAGAAGGGCGTCTCAGGAGGGCTCGGGTTTGCTTTATTTTATTTT
GTCATTGCGGGGTATATGGAGAGTCA

57/246

FIGURE 56

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA131590
><subunit 1 of 1, 497 aa, 1 stop
><MW: 55906, pI: 8.43, NX(S/T): 4
MSCVLGGVIPLLFLVCGSQGYLLPNVTILLELLSKYQHNESHSRVRAIPREDKEEIL
MLHNKLRGQVQPQASNMEYMTWDDELEKSAAWASQCIWEHGPTSLVSIGQNLGAHWGR
YRSPGFHVQSWYDEVKDYTYPYPSECNPWCPERCSPGMCTHYTQIVWATTNKIGCAVNTC
RKMTVWGEVWENAVYFVCNYSPKGNWIGEAPYKNGRPCSECPPSYGGSCRNNLCYREETY
TPKPETDEMNEVETAPIPEENHVWLQPRVMRPTPKKTSAVNYMTQVVRCDTKMKDRCKG
STCNRYQCPAGCLNHAKIFGSLFYESSSSICRAAIHYGILDDKGGLVDITRNGKVPFFV
KSERHGVQSLSKYKPSSSFMVSKVKVQDLCYTTVAQLCPFEKPATHCPRIHCPAHCCKDE
PSYWAPVFGTNIYADTSSICKTAVHAGVISNESGGDVDVMPVDKKTYVGSLRNGVQSES
LGTPRDGKAFRIFAVRQ
```

Important features of the protein:

Signal peptide:

Amino acids 1-22

N-glycosylation sites:

Amino acids 27-31; 41-45; 451-455

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

Amino acids 181-185; 276-280; 464-468

Tyrosine kinase phosphorylation site:

Amino acids 385-393

N-myristoylation sites:

Amino acids 111-117; 115-121; 174-180; 204-210; 227-233; 300-306;
447-453; 470-476

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 2:

Amino acids 195-207

SCP-like extracellular protein:

Amino acids 56-208

58/246

FIGURE 57

GCACGAGGCCAACACACAGCAGCCTCAACATGAAGGTGGTTATGGTCCTCCTGCTGCTGCC
TCCCCCTTACTGCTATGCAGGTTCTGGTTGCCTCTGGAGAGCGTCGTGGAAAAGACC
ATCGATCCATCGGTTCTGTGGAGGAATACAAAGCAGATCTCAGAGGTTATCGACACTGA
GCAAACCGAAGCAGCTGTAGAGGAGTTCAAGGAGTGCTCCTCAGCCAGAGCAATGAGACTC
TGGCCAACCTCCGAGTCATGGTCATACGATATATGACAGCCTTACTGTGCTGCGTATTAA
CTGTCACAAGAACCTTGGCTCAGAGGAATCCAGACGATGCTCACAACCCACTGTGGACTGG
CAGAAATCTCAACTTTCCCTTGACTTTCCCCTTGATCAGTAATATGGAAGACGTTGTTG
AAACCTGAAGTATAAGTTAATTAAATAACCCACTGCAAGAAAAAAAAAAAAAA

59/246

FIGURE 58

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA135173
><subunit 1 of 1, 93 aa, 1 stop
><MW: 10456, pI: 4.37, NX(S/T): 1
MKVVMVLLAALPLYCYAGSGCVLLESVVEKTIDPSVSVEEYKADLQRFIDTEQTEAAVE
EFKECFLSQSNETLANFRVMVHTIYDSLYCAAY
```

Important features of the protein:

Signal peptide:

Amino acids 1-18

Prokaryotic membrane lipoprotein lipid attachment site:

Amino acids 12-23

FIGURE 59A

CAAGTCGTTGAGGCTGCCAGGCAGTCAGGTCTCTGGACCTCGCCTGACTCGGCTGGC
 TGTGCCTGAAATTGACCCAGCTCCACCATACTCCTGATT**ATGAGAAAACAAGGAGTAAGCT**
 CAAAGCGGCTGCAATCTCCGGCCGAGCCAGTCTAAGGGGCGGCGGGGGCTCCCTGCC
 CGGGAGCCGGAGGTAGAGGAGGGTGGAAAAGTCGGTCCTAGGCAGGGAAACTGCCAAG
 GGGCGCCTGGAGGTCTCCCCGGGGAGGATCAAAGTCTGAAAGAGCGAAAAGGCTGGAGC
 TAGAGGTGGTGGCAAGACCTTCTCTCGGCCCCCTCCAGTCGTCCGTAATTCCCTGGCG
 CAGCTCCGGAAAAGGTGCAGGAACTGCAGGCGCCGGTTCTCCAGCAGAACCACTCTCGG
 CATCGCTGTCTTGTGGCAATTTCATGGTTACATTAGTAACACTTTGAAAATGATC
 GTCATTCTCTCACCTCATCTTGAACGGGAGATGACTTTCGCACTGAAATGGACTT
 TATTATTCAACTCAAGACCATTATTGAAGCACCTCGTTTTGAAAGGACTGTGGATGAT
 TATGAATGACAGGCTACTGAATATCCTTATAATTAAATGCAATAAAACGCTCCATCTT
 ATCCAGAGGTAATCATAGCCTCTGGTATTGCACATTCACTGGAAATAATGAATTATTGGA
 CTAGAAACTAAGACCTGCTGGAATGTCACCAGAATAGAACCTCTTAATGAAGTTCAAAGCTG
 TGAAGGATTGGAGATCCTGCTTGCTTTATGTTGGTGTAACTTTTATTAAATGGACTAA
 TGATGGGATTGTTCTCATGTATGGAGCATACCTGAGTGGGACTCAACTGGGAGGTCTTATT
 ACAGTACTGTGCTTCTTTCAACCATGGAGAGGCCACCCGTGTGATGTGGACACCACCTCT
 CCGTGAAAGTTTCTATCCTCTTGACTTCAGATGTTGATTTAACTTGTGATTCTCA
 GGACCTCAAGCAATGATAGAAGGCCCTCATTGCACTCTGCTTCCAATGTTGCTTTATG
 CTTCCCTGGCAATTGCTCAGTTTACACAGATAGCATCATTATTCCATGTA
 TGTGTGGGATACATTGAACCAAGCAAATTCAAGAAGATCATTATATGAACATGATTCAGTT
 ACCCTTAGTTCAATTGATGTTGAAATTCAATGTAATTCTTATTATTCTTCATC
 TTTGTAATGACGTGGCAATAATTCTAAAGAGAAATTCAAAAACGGGAGTATCTA
 AACTCAACTTTGGCTAATTCAAGGTAGTGCCTGGTGTGGAACAATCATTGAAATT
 CTGACATCTAAATCTTAGGCGTTCAGACCACATTGCCTGAGTGTATCTATAGCAGCCAG
 AATCTTAAGGTATACAGATTGATACTTAAATATACCTGTGCTCCGAATTGACTTCA
 TGGAAAAGCGACTCCGCTGAGATACACAAAGACATTATTGCTCCAGTTGTTATGGTGA
 ACATTTTATCTTAAAGACTGTTGATATTCAATGTTAGCTACAAACATTAA
 TCTAAGAAAACAGCTCCTGAAACACAGTGAGCTGGCTTTCACACATTGCAATTGTTAGTGT
 TTACTGCCCTTGCCATTAAATTGAGGCTAAAGATGTTTGACACCGCACATGTGTGTT
 ATGGCTCCTTGATATGCTCTGACAGCTTGGCTGGCTTTGCAAGGTTGTTGTTGA
 GAAGGTTATCTTGGCATTAAACAGTGTCAATACAAGGTATGCAAACCTCCGTAATC
 AATGGAGCATAATAGGAGAATTAAATTGCTCAGGAAGAACATTACAGTGGATCAA
 TACAGTACACACATCAGATGCTCTTGCAAGGTGCCATGCCATACAATGGCAAGCATCAAGCT
 GTCTACACTCATCCCATTGTGAATCATCCACATTACGAAGATGCAAGACTTGAGGGCTCGGA
 CAAAATAGTTATTCTACATATAGCAGAAATCTGCCAAGAAGTAAGAGATAAATTGTTG
 GAGTTACATGTGAATTATTGTTAGAAGAGGCATGGTGTGTTGAGAAGTAAGCCTGG
 TTGCACTGTTGAAATCTGGGATGTGGAAGACCCCTCCAATGCAAGCTAACCCCTCCATT
 GTAGCGCTGCTCGAAGACGCCAGGCCTACTTCACCACAGTATTTCAGAATAGTGTGTAC
 AGAGTATTAAAGGTTA**ACTTGAGAAGGATACTACCCATTAACTATGGCACAATGCCGTGT**
 CAAAACAATCACCTTGGCTTATTCACTTAATAAAATCACAAGCTTAATAACAGACA
 CTTAAAAATAAGATAAAATGGATTGAAATTCTGATTACTAAAGGTAATTACTTT
 CTGTTCATGAAATGTCAGCCTTATTAAGCTGTCACTAAAGTTATTAAATCATTCACTGTCAT
 ACTGCATAAACAAATGTTCAAGATTAAAGAGAAATGTATATAAGAACMATGATT
 TTAATAATCAGGGTATGTAAGCCTTTCATCCAACTAGGTGAATTGCTTCAGATTCT
 CTAGTACCAAGAGGGTACCTCCTCAAACCTTGTGAACCACTTAAGGCAGAAGAATGCAAGCTC
 TGAAATGACATCCTAAATGCTGATACTGGTCACAGCCTTACCTCTGTGAGGAAATTG
 TAACAGTGTGTTAAGGTGTTTATTACAGGCCCTAAGAAAGATCTCTAATAACCT
 TTTAATAACTTTTTAATAATTCAAGGTGAAGTGTGTTAAAACACTTGTGTTGAAT
 GTTTGAATCTCTTGAGATGTGTTACCCACTAGATACTATTGCCACTGGTAGTTCTC
 CATCTAACGCTCAAGAGGTTATTCACTCTCTTAGATTCCAGTGGCTTCTTAAACATCC
 AGGTAAAACAGAAACTGCTATGGTATACAACCAAGTTTGGGTTAAACATAATCAGAAAAG

61/246

FIGURE 59B

AAAATCCAGTTAAATTATGAAGTGAGATTTCAGATCCTAGATCTTGAATAAAGGAAAGGT
CTTTCATCTTGATGGCCCCAAAGCTTGGTGCATGGCTTTATTTCTGGCCACTATCTC
TTAAATAATATATTTAAGCCCTCATTATTTGGTTTGGGTGAGGAAAGTCATGTTT
CTAAGTCCTCTCCCCTAATAAAACCTACCCAAACAATAGTGCCTTGAAAGTGGTAGTTATCT
TGAAGATACTCTGCCAAATGCAAAGATAAACATCTTTGTCTGCCTTATAAATATGAA
TATGCCAGATCTATAGTATTTAATGTGCATCTACTTAAATGAGTCATCTGGGTTTTA
TAATTCCCTTATGTTCTGCCCTCTACACTTGAAATAACAAAATGCCTTAATTTATGGAT
TAGTTCTCTTATAGTAGACAGGCAGCTATATGCAGCAAAACCAATAAAGTTATTTCAACT
TTCATAGTTGTAAAATATCTTATACCAGAATACAAAACAGCTAAGAAAACATGCCACATTTAT
TTTAGCATTCAAATAATTGTTTGGTGTAAAGCACAGGATAAAAAGGAGAGCGTCAA
GAAAAGAGACATAACACCTAACATTCAAAAAATTAAACAAAGTATATTGGATGATGTTT
TACAGGAAATATTTAAATAAGTTGGTAGAACTTTAAAATGGTACTGTATTAGCTAATAAA
ATATTCAAGTACAAATATATGTTGGATTATGCATTAAAAACTAATAAAATTATTCAC
TTA

62/246

FIGURE 60

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA138039
><subunit 1 of 1, 758 aa, 1 stop
><MW: 87354, pI: 9.36, NX(S/T): 1
MRKQGVSSKRLQSSGRSOSKGRRGASLAREPEVEEVEKSVLGGGKLPRGAWRSSPGRIQ
SLKERKGLELEVVAKTFLLGPFQFVRNSLAQLREKVQELQARRFSSRTTLGIAVFVAILH
WLHLVTLFENDRHFSLSSLEREMTFRTEMGLYYSYFKTIIEAPSFLEGLWMIMNDRLTE
YPLIINAIRKFHLYPEVIIASWYCTFMGIMNLFGLETKTCWNVTRIEPLNEVQSCEGLGD
PACFYVGVIIFILNGLMMGLFFMYGAYLSGTQLGGLITVLCFFFNHGEATRVMWTPPLRES
FSYPFLVLQMCILTLILRTSSNDRRPFIALCLSNVAFMLPWQFAQFILFTQIASLFPMYV
VGYIEPSKFQKIIYMMNISMVTLSFILMFGNSMYLSSYYSSSSLLMTWAIILKRNEIQKLGV
SKLNFWLIQGSAAWCCTIILKFLTSKILGVSDHIRLSDLIAARILRYTDFDTLIYTCAPE
FDFMEKATPLRYTKTLLPVVMVITCFIFKKTVRDISYVLATNIYLRKQLLEHSELAFHT
LQLLVFTALAILIMRLKMFLTPHMCVMASLICSRQLFGWLFRVRFEKVIFGILTVMMSIQ
GYANLRNQWSIIGEFNNLPQEELLQWIKYSTTSDAVFAGAMPTMASIKLSTLHPIVNHPH
YEDADLRARTKIVYSTYSRKSAKEVRDKLLELHVNYYVLEEAWCVVRTKPGCSMLEIWDV
EDPSNAANPPLCSVLLLEDARPYFTTVFQNSVYRVLKVN
```

Important features of the protein:**Transmembrane domain:**

Amino acids 109-124; 197-213; 241-260; 266-283; 302-315; 336-356;
376-391; 430-450; 495-509; 541-560; 584-599; 634-647

N-glycosylation site:

Amino acids 222-226

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 102-106

Tyrosine kinase phosphorylation site:

Amino acids 511-519

N-myristoylation sites:

Amino acids 24-30; 50-56; 151-157; 254-260; 264-270; 269-275;
273-279; 639-645

Amidation site:

Amino acids 20-24

FIGURE 61

GGCGCGGCCACATCCTTAAATATGGTCTTCTGGGCGCGCGACAATGTGAGGAGTGGG
GTGGAGCGTGTGGTGTGGCTGCAGCCTGGCAAGAGCCGCGGACCATGAGCTGAG
TAAGTTCTGGAGGGATCCTGCCTCTGGAGCCTCGCAGCCAGGCAGCTGTGAACGTGAGC
TAGAGTGAAGCAGAAATCTAGGAAGATGAGCTCCAAGATGGTCATAAGTGAACCAGGACTGA
ATTGGGATATTCCCCAAAATGGCCTTAAGACATTTCTCGAGAAAATTATAAGAT
CATTCATGGCTCCAAGTTAAAAGAACATCGTGTATTCACAGACGTATAGGAGAAA
TTTGAATGCCTCAGCAAGTTCTGTAGAAAATGAGCCGGCAGTTAGTCAGCAACTCAAGCAA
AGGAAAAAGTTAAAACCACAATTGGAATGGTCTTCTCCTAACCAAGAGTTCTTAC
CGTTCTCGTTCTCACAGAGAGCAGAGGAGTTATGTGGACTTGGTTAAATACGC
AAAGATTCTGCAAATTCAAAGCTTGAATAAATAAAATGACTACTTGCAGTACTTGG
ATATGAAAAAAACATGTGAACGAAGAAGTTACTGAGTTCTAAAGTTTGAGAATTCTGCA
AAGAAATGTGCGCAGGATTATAATATGCTTCTGATGATGCCGTCTTCACAGAGAAAAT
TTTAAGAGCTTGCATTGAACAAGTGAAGAAAGTATTCAAAGTATTCTATACTCTCCACGAGGTCA
CCAGCTTAATGGGATTCTTCCCATTAGAGTAGAGATGGGATTAAGTTAGAAAAAACTCTT
CTCGCATTGGGCAGTGTAAAATATGTGAAAACAGTATTCCCTCAATGCCTATAAGTTGCAG
CTGTCAAAGGACGATAGCTACCATGAAACAGTCAGAACAAACAGCTGAAGCTATGCATTA
TGATATTAGTAAAGATCCAAATGCAGAGAAGCTTGTGTTCCAGATATCACCTCAGATAGCTC
TAACTAGTCAGTCATTATTACCTTATTAAATAATCATGGACCAACGTACAAGGAACAGTGG
GAAATTCCAGTGTGTATTCAAGTAATACCTGTTGCAGGTTCAAAACAGTTAAAGTAATATA
TATAATTCAACCACCTCCCCAAAAGAAAATGACTATGAGAGAGAGAAATCAAATTTCATG
AAGTCCATTAAAATTATGATGTCCAAAACACATCTGTTCCAGTCTGCAGTCTTATG
GACAAACCTGAAGAGTTATATCTGAAATGGACATGTCCTGTGAAGTCAACGAGTGCCGAAA
AATTGAGAGTCTGAAAACCTGTATTGGATTTGATGATGTCAAGAACTGAAACTT
TTGGAGTAACCACCAACCAAGTATCAAATCACCAAGTCCAGCAAGTACTTCCACAGTACCT
AACATGACAGATGCTCTACAGCCCCAAAGCAGGAACACTAACACTGTGGCACCAAGTGCACC
AGACATTCTGCTAATTCTAGAAGTTATCTCAGATTCTGATGGAACAATTGCAAAGGAGA
AACAGCTGGTCACTGGTATGGATGGTGGCCCTGAGGAATGCAAAATAAGATGATCAGGGA
TTGGAATCATGTGAAAAGGTATCAAATTCTGACAAGCCTTGTATACAAGATAGTGA
AACATCTGATGCCTTACAGTTAGAAAATTCTCAGGAAATTGAAACTCTAATAAAAATGATA
TGACTATAGATATACTACATGCTGATGGTGAAGACCTAATGTTCTAGAAAACCTAGACAAC
TCAAAGAAAAGACTGTTGGATCAGAACAGCAGAAAAACTGAAGATACAGTTCTGCAGCAG
TGATACAGATGAGGAGTGTAACTCATGATACAGAATGAAAAAAACAGTTATAACAGTGT
TTTAATTAGATAAGTTGAGGGAAAATAATCAGTAGGCAAGAGGAACATTTCCTGTAGT
AGCTAGAGTGCCTTGAAAAAATGTGTTGGCTATGTGAAGGAATATTCAACTAAAATGGAAT
GGTATGCTTTCACCCCTAAAGTTGAGGAGGATCTGATATGTTAACATTATCATGGCA
GGGAAATATAAGAAGAAAATATTTACATTAACCTTTCTAAAATTGAAATAGA
AAAATAATTGGTTTTATCAAGAACACACTATCGTTATGTATTGTTAGTTATATTG
CCAGTCTGTTGCGACTGACTCAAAAAGTTAAATGTTGCCACTGCTGAAGATGATTGAGCA
TCGCAAACCTTGTGACCCATTGACAGTTTATATACTCCTTAAAATGATGAATG
TTACAGGTTAATAAGTTAATACCTTAA

64/246

FIGURE 62

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA139540
><subunit 1 of 1, 592 aa, 1 stop
><MW: 66453, pI: 5.42, NX(S/T): 3
MSSKMWVISEPGLNWDISPKNGLKTFFSRENYKDHSMAPSLKELRVLSNRRIGENLNASAS
SVENEPAVSSATQAKEKVTTIGMVLPLPKPRVPYPRFSRFSQREQRSYVDLLVKYAKIPA
NSKAVGINKNDYLQYLDMKKHVNEEVTEFLKFLQNSAKKCAQDYNMLSDDARLFTEKILR
ACIEQVKKYSEFYTLHEVTSLMGFFPFRVEMGLKLEKTLLALGSVKYVKTVFPSMPIKLQ
LSKDDIATIETSEQTAEAMHYDISKDPNAEKLVSRYHPQIALTSQSLFTLLNNHGPTYKE
QWEIPVCIQVIPVAGSKPVKVIYINSPLPQKKMTRERNQIFHEVPLKFMMSKNTSVPVS
AVFMDKPEEFISEMDMSCEVNECRKIESLENLYLDFDDDVTLETFGVTTTKVSKSPSPA
STSTVPNMTDAPTAPKAGTTVAPSAPDISANSRSLSQILMEQLQKEKQLVTGMDGGPEE
CKNKDDQGFESCEKVSNSDKPLIQDSDLKTSDALQLENSQEIETSNKNDMTIDILHADGE
RPNVLENLDNSKEKTVGSEAAKTEDTVLCSSDTDEECLIIDTECKKTSYNSV
```

Important features of the protein:

N-glycosylation sites:

Amino acids 56-60; 354-358; 427-431

cAMP- and cGMP-dependent protein kinase phosphorylation sites:
Amino acids 187-191; 331-335; 585-589

N-myristoylation sites:

Amino acids 126-132; 407-413; 557-563

65/246

FIGURE 63

TTTTAACTTGAACCTCCAAGGCCACGTGCGTCTCCTGGCTCCTGCACGGACTGTGTGACTG
TCCCCGACAGCTTCCTGTCTCGTCTCATGAGGGGTCCAGCACATGGCATTCTGGGTGGCA
CCTGAAGTCCACCTCTATGAGACCCCTGGGAGCGTGACGGGGCCTGGC**ATGGGTGGCCG**
AGGCCCTCTGTCCCAGGTCACTGGTGTGGTCGGCCAGGCCCTCTGTCCCACATCACCTG
TGTGGTCGGCCAGGCCCTCTGTCCCAGGTCAACCGGTGTGGTCGGCCAGGCCCTCTGTG
CAGGTCTCCTGTCCCAGGTCACTGGTGTGGTCGGCCAGGCCCTCTGTCCCAGGTCAACCTG
TGTGGTCGGCCAGGCCCTCTGTACCATGTCAGTGTGAGGGGCTGGCTCTGGAAGAGGG
CAGGGACTTGGCATTGGTGGGGCAGGGTCCAAGGTGTGGCTGTCAAGCAGGAAGGGCAG
GTGGCATGGTCCAGGCAGGACTCAGGGTGGGGACTGCTGGAGACTGTCCGGAGGCC
CCTCCAGGGCACCTGCCATTGCCATTGTCGTCATGCCATCTGGTCCCCTTCAGGGAAC
AAGAGGAGGATCAGATGCTGGGGACATGAT**TGAGAAGCTGGGTGACTGGCCGGGATGCT**
GAGGGCTGGCTGGCTGGCTGGGTGGGCCGGGATGCTGAGTGCTGGGCTGGCTGGCTGGGT
GGACCGGGCCTCCAGCTGGGGTGGGGGGGGGGGTATGGGTCCCCCTCAGCCTTGG
TGACAGGACAGGCAGGTCACCTGAGGGTGAGAGCTCCCTCCGCCCTAAGAGAGCCAGG
GGCAGCTGGTGACCGTGTGGTCATGGTGGGACAGCCCTCCGGGGCACCCAGTCGGGGCAG
GTCTCACGTGGAGGGCACAGGGCTCTGCAGGCTCGGAGGCCAGGGCGGATTGTGGCC
AGTGGAAAGGAAAGGATGTTCTGGCAGGGGACTGTGTGGCCACGGCTGTGGCTGGGG
CGTTGAGCACGGCCTCACTGTCCACCTGTCCCCTAGGCCTCCAGAGGAAGAAGTCCAAGTTC
CGCTTGTCCAAGATCTGGTACCAAAAAGCAAAGCAGCCCTCCAGTAGTACGCCAGTAGG
GCCGTGGCTCGGCCGGACCTGGCATCCGACTTGGACTCGGGCCATGGCTGGCCCG
ACCCGGAACCCGGACTTGTACTCGGGGCCGTGGCTCGGCCGGACCCGGATTGGACTTGG
GACTCGGGAAAGGGCCTCTGTCCCCTACAAGGGGATGTGGACAGCAGGGACCTGCGCTACCG
TCTGTGGTCTCAATAAGAAACCGACCACATGGCCCCGGAAAAAAAAAAAAAAACA
AAAAAAAAAAAAAAACA

66/246

FIGURE 64

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA139602
><subunit 1 of 1, 159 aa, 1 stop
><MW: 15900, pI: 8.07, NX(S/T): 0
MGRPRPFCPRSLVWSAQALLSHITCVVGGPPVPGHRCGRPRPSCPAGPVQVTGVVGP
GP
SVPGHLCGRPRALLYHVTVEGLALEEGRDLALVGAGFQGVACQQEGAGGMGPGGTQGWGA
TAGDCPEAPPGHLAIAIAVAHGLVPFQGTRGGSDAAGHD
```

Important features of the protein:

Signal peptide:

Amino acids 1-25

N-myristoylation sites:

Amino acids 109-115;113-119;119-125;148-154;151-157;152-158

67/246

FIGURE 65

GGCGACCACCGCCGCCTCCTCACCTGGCATTGGTGCAGCCGTTCCGGCGAGAGAAG
GCAGGGCGCTCCGCGGGCTGGGTCCGGCACTCTGACCCCCCTTGTAACCACCGCGGCCG
ATGGTGGGCTCCGCGGGCTGGGTCCGGCACTCTGACCCCCCTTGTAACCACCGCGGCCG
CACCCAGGGAGTTGAGCAACGAAGTTGGTGCACCTGCCCCGCTCCCAGGCAGTTGCTGTG
GGGCTTCACGGCTGCTGGAAGGGCATGGCTGTTGCTCCCATCACTGGCGCCAGCTCTCA
AAGCTACGTTACAGCAACGCAGTAGGGACTTCGTGGCAGGCTTTTAAGAGCTGAAAG
AAGGGCAGGGAGGGTTACGTCT**AGGGT**GATGATTCTCACCAAGACAGCGAAGTATCTATT
GGGAAACTCCAGGTGACCGCACCTCCTCCGACAGTTGCCCCGGGCAAGTTACCAGCTG
CGTCAGAAAGCAGGTTGCAAATCCTGGAGAACGGCCTGAGCTAAGGACTGGGTCAGGA
GGGTTTAAACTCATCTGATTCTTGCATCATATCTCTGAAAGTTTATTTCCC
CAATATTTCTGAGTTGCTATATCCAATGAAAACAATGCTGATGTAGAGGTCCACCAGCCA
ATGCTTATTGGAAGTCAACGAATGAGACCGAGGGTGGCCCATAATCAATCTGGCACGCGG
GAATGTGAAACCTCTTCAAGGTCTGGCGAGTCCCTAGAGTTACGCAGATGAAGGACATTGG
CCCTCGAGAATCTCACACCAGCAAAGAAGAGCACACGAAGCGAAACTACTTATGATCATT
GTGGCTTGGCAAGTTGTTGAGCTCCAGCAACAATTCTCACCTGGAGTGCAGCAATA
AATGATACTGGTGCAGGGCAGCTAATAAGCTTCTGAATAATATGCAAAGTACTTGGC
ACCATGAGCAGAACTCAGTATACCGTCACTGAAGAAATAGCTTATTAATGATTACACTTT
CATATGTGCAAGTAAAAGTTGACTTTAGGGAGGCCTCACCTACGGAATGTCTTTAA
ATTCTTTTAATTATACTTAAAGTTCTGGGATACATGTGCAGAACGTGCAGGTTGTTAC
ACAGGTATACATGTGCCATGGGGTTGCAGCACCCATCAACCCTCATCTAGGTTAAAGC
TCCGCATGCATTAGTTATGTCTTAATGCTCTCCCTCCCTGTCCCCCACCCCCAACAG
GCCTCAGGGTGTGATGTTCCCTCCCTGGGTCCATATGTTCTCATTGTTCACTCCACTTA
TGATGAGAACATGCAGTGTGGTTCTGTTCTGTGTTAGTTGCTGAGAATGATGGTT
CCAGCATCATCCACGTCCCTGCAAAGGACATGAATTCAATTCTTTATGGCTGCATGGTAT
TCCATGGTGTATATGTGCCACATTTCTCATCCAGTCTATCATGATGGGCACTTGGGTTG
GTTCCAAGACTTGTATTGTGAACAGTGCTGCAATAAACATACGTTGTATGTGCAAAAAA
AAAAAAAAAAAAAAA

68/246

FIGURE 66

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA139632
><subunit 1 of 1, 90 aa, 1 stop
><MW: 9586, pI: 12.18, NX(S/T): 0
MVGSARLGPELLTPVTTAAGTQGVRATKLVTCAPRQFAVGAFTAAAGRAWLFVPSLGAS
FSKLRQQRSRDFRGRLFLRAERRAGGFTS
```

Important features of the protein:**Signal peptide:**

Amino acids 1-24

N-myristoylation sites:

Amino acids 24-30; 42-48; 58-64

69/246

FIGURE 67

CATGTCTAGACTGGGAGCCCTGGGTGGTGCCCGTGCCTGGACTGGACTGTTGCTGGGTACCG
CCGCCGGCCTTGGATCCTGTGCCTCCTTACAGCCAGCGATGGAAACGGACCCAGCGTCAT
GGCGCAGCCAGAGCCTGCCAACCTCCCTGGACTATACGCAGACTTCAGATCCCGGACGCCA
CGTGATGCTCCTGCAGGCTGTCCAGGTGGGGCTGGAGATGCCTCAGTGCTGCCAGCCTTC
CACGGGAAGGACAGGAGAAGGTGCTGGACCGCCTGGACTTGTGCTGACCAGCCTGTGGCG
CTGCGGCGGGAGGTGGAGGGAGCTGAGAAGCAGCCTGCGAGGGCTGCGGGGGAGATTGTTGG
GGAGGTCCGATGCCACATGGAAGAGAACAGAGAGTGGCTCGGCAGGCTGGAGGTTCCGTTG
TCCGGGAGAGGAGTGACTCCACTGGCTCCAGCTGTCTACTTCACGCCCTCCTGGGAGCC
ACGTTACAGATGCTGAGGTGAGAGTGAAGGGGTTACACAACAGCCAATGCGGAGTCTGACAATGA
GCGGGACTCTGACAAAGAAAGTGAGGACGGGGAAAGATGAAGTGAGCTGTGAGACTGTGAAGA
TGGGGAGAAAGGATTCTCTTGACTTGGAGGAAGAGGGCAGCTCAGGTGCCTCCAGTGCCTG
GAGGCTGGAGGTTCCCTCAGGCTTGGAGGATGTGCTGCCCTCCTGCAGCAGGCCAGAGCT
GCACAGGGGTGATGAGCAAGGCAAGCGGGAGGGCTTCCAGCTGCTGCTCAACAACAAGCTGG
TGTATGGAAGCCGGCAGGACTTCTCTGGCGCTGGCCCAGCCTACAGTGACATGTGTGAG
CTCACTGAGGAGGTGAGCGAGAAGATCATATGCCCTAGATGAAAAGAAGAAGCAGAGGC
TGCTCTGGAGAAGGGGATGAGAGTGACTGTGACTGTGACCTGTGGTATGCCGTGCTTGTGGTC
AGCTGGCTGAGCATGAGAGCATCCAGAGGCCATCCAGAGTGCTTAAAGGAGC
GTGGACAAAGCCATTGCTCTCAGCCAGAAAACCCATGGCTCACCTCTTGGCAGGTG
GTGCTATCAGGTCTCACCTGAGCTGGCTAGAAAAAAACTGCTACAGCCTGCTTAAAGGAGC
GCCCTCTCAGTGCCACTGTGGAAGATGCCCTCAGAGCTTCTAAAGGAGC
CCAGGATTTCCAAAGCAGGAAGGTATATATTCAGTGCCTACAGAGAACTAGGGAAAAA
CTCTGAAGCTAGATGGTGGATGAAGTGGCCCTGGAGCTGCCAGATGTCACGAAGGAGGATT
TGGCTATCCAGAAGGACCTGGAAGAACTGGAGTCATTAACTGAGACTAACCACGTTCACT
GCCCTCATGACTGATGCCACTATTAAGGTGGGGGGGGAGGCTTTCTTAGAC
CTTGCTGAGATCAGGAAACCACACAAATCTGCTCTGGTCTGACTGCTACCCACTACCAC
TCCCCATTAGTTAATTATTCTAACCTCTAACCTAACTGAGAATTGGGGCAGTACTCATGGC
TTCCGTTCTGTTCTCTCCCTGAGTAATCTTAAAGGAGC
CCAGGATTACACATGGTAGAGCCTGCAAGACCTGAGACCTCCAATTGCTGGTGGAGTGG
TGAACCTCAAAGCTATAGGAACAAAGCACATAACTGTCACTTAACTTTTCACTGACTA
ATAGGACTCAGTACATATAGTCTTAAGATCATAACCTTACCTACCAAGGTAAAAGAGGGATCA
GAGTGGCCCACAGACATTGCTTCTTACCTATCATGTGAATTCTACCTGTATTCCCTGGG
CTGGACCACTTGATAACTTCCAGTGTCTGGCAGCTTGGAAATGACAGCAGTGGTATGGGG
TTTATGATGCTATAAAACAATGCTGAAAAGTTGCCTAGAATATTTGTTACAAACTTGA
AATAAACCAAATTGATGTT

70/246

FIGURE 68

```
>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA139686
><subunit 1 of 1, 470 aa, 1 stop
><MW: 52118, pI: 5.06, NX(S/T): 0
MSRLGALGGARAGLGLLLGTAAGLGFLCLLYSQRWKRTQRHGRSQSLPNSLDYTQTSDPG
RHVMLLRAVPGGAGDASVLPSPREGQEKVLDRLDFVLTSIVALRREVEELRSSLRGLAG
EIVGEVRCHMEENQRVARRRRFPFVRERSDSTGSSSVYFTASSGATFTDAESEGGYTTAN
AESDNERDSDKESEDGEDEVSCETVKMGRKDSLDLEEEAASGASSALEAGGSSGLEDVLP
LLQQADELHRGDEQGKREGFQLLLNNKLVYGSRQDFLWRLARAYSDMCELTEEVSEKKSY
ALDGKEEAEAALEKGDESADCHLWYAVILCGQLAEHESIQRIQSGFSFKEHVDKAIALQP
ENPMAHFLLGRWCYQVSHLSWLEKKTATALLESPLSATVEDALQSFILKAELQPGFSKAG
RVYISKCYRELGKNSEARWWMKLALELPDVTKEDLAIQKDLEEVILRD
```

Important features of the protein:**Signal peptide:**

Amino acids 1-32

cAMP- and cGMP-dependent protein kinase phosphorylation site:
Amino acids 209-213**N-myristoylation sites:**Amino acids 5-11;8-14;9-15;15-21;19-25;72-78;164-170;
174-180;222-228;230-236**Amidation sites:**

Amino acids 207-211;254-258

Cell attachment sequence:

Amino acids 250-253

FIGURE 69

CCCCACCGCGTCCGAAACACTTTAACCTGACCAGCTAAATGGATAAAACCTAGCCTGCATAGCT
TTTAAACTGGGTCTCATACAGCACAGGAGGCCTACTGCTTCAAGAACTGAAAATCCAGAG
GATGAATTGCTTATCTGGATGCCAAAAGCCAGCACAATAAGGAATGCCAGTTGTATGG
GGCTACTAGCTCACATGCCGATCAGAATGGTGTGAATGACAGCCGACTGTGTATGAAGG
TGGTGGTGGTTCCGCACAAGAGACCAATAAGAAGAAAGCTGAGAGAGGGGGGAAACGTTTT
GGATGACAAGGATGGTTCCATTAAATTACGCAGCTGAAAGGCATGAGTGTGGTGTGGT
GCTACTCCTACACTGCTGCTGTATGCTCACGGGTGCTCAGAGAGCTTGCCTAAAGAACT
GCAGATGTGATGGCAAATTGTGTACTGTGAGTCTCATGCTTCGCAGATATCCCTGAGAAC
ATTTCTGGAGGGTCACAAGGCTTATCATTAAAGGTTAACAGCATTAGCTCAAATCCAA
TCAGTTGCCGGCTTAACCAGCTTATATGGCTTATCTGACCATAATTACATTAGCTCAGTG
GATGAAGATGCATTCAAGGATCGTAGACTGAAAGAATTAAATTCTAAGCTCAAACAAAAT
TACTTATCTGCACAATAAAACATTCAACCCAGTTCCAATCTCCGCAATCTGGACCTCTCCT
ACAATAAGCTTCAGACATTGCAATCTGAACAAATTAAAGGCTTCGAAACTCATCTTTG
CACTTGAGATCTAACTCACTAAAGACTGTGCCATAAGAGTTTCAAGACTGTCGGAATCT
TGATTTTGGATTGGTTACAATCGTCTCGAAGCTTGTCCGAAATGCATTGCTGGCC
TCTTGAAGTTAAAGGAGCTCCACCTGGAGCACACCAGTTCCAAGATCAACTTGCTCAT
TTTCCACGTCTCTCAACCTCCGCTCAATTACTTACAATGGAACAGGATTGCTCCATTAG
CCAAGGTTTGACATGGACTTGGAGTTCTACACAACCTGGATTATCAGGGAAATGACATCC
AAGGAATTGAGCCGGGCACATTAAATGCCTCCCCAATTACAAAAATTGAATTGGATTCC
AACAAAGCTACCAATATCTCACAGGAAACTGTCAATCGTGGATATCATTAAATATCCATCAC
ATTGTCGAAATATGTGGATGCACTGGAGCATTGTCTTATTGCTTATTGCTTAAGA
ATTCAAAAGGAATAAGGAAAGCACCAGTATGTGGGGACCTAAGCACATCCAGGGTGA
AAGGTTAGTGATGCAGTGGAAACATATAATATCTGTTCTGAAGTCCAGGGTCAACACAGA
AAGATCACACCTGGTGCCCAAACCTCCCCAGAAACCTCTGATTATCCCTAGACCTACCATCT
TCAAACCTGACGTACCCAAATCCACCTTGAACACCAAGCCCTCCCCAGGGTTTCAAGATT
CCTGGCGCAGAGCAAGAGTATGAGCATGTTCATTCACAAAATTATTGCCGGAGTGTGG
TCTCTTCTCTCAGTGGCCATGATCCTTGGTGTATGTCATTGCTTGGAAACGCTACCCAG
CCAGCATGAAACAACCTCCAGCAACACTCTCTTATGAAGAGGCGGGAAAAAGGCCAGAGAG
TCTGAAAGACAAATGAATTCCCTTACAGGAGTATTATGTGGACTACAAGCTACAAACTC
TGAGACCATGGATATATCGGTTAATGGATCTGGCCCTGCACATATACCATCTGGCTCCA
GGGAATGTGAGATGCCACACCACATGAAGCCCTGCCATATTACAGCTATGACCAGCCTGTG
ATCGGGTACTGCCAGGCCACCAAGCCACTCCATGTCACCAAGGGCTATGAGACAGTGTCTCC
AGAGCAGGACGAAAGCCCCGGCTGGAGCTGGGGAGACCAAGCTCATGCCACCATCG
CCAGGTGGCAGCACCGGCATCTACCTAGAGAGAATTGCAAACTAACGCTGAAGCCAACCTC
CTCACTGGGGAGCTCCATGGGGGGAGGGAGGGCCTCATCTTAAAGGAGAATGGGTGTCCA
CAATCGCGCAATCGAGCAAGCTCATCGTCCGTTAAAACATTATGGCATAGGGAAAAAAA
AAAAAAAAAAAAAA

72/246

FIGURE 70

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA142392
><subunit 1 of 1, 590 aa, 1 stop
><MW: 67217, pI: 9.26, NX(S/T): 4
MGFHЛИTQLKGMSVVLVLLPTLLLVMLTGAQRACPKNCRCDGKIVYCESHAFADIPENIS
GGSQGLSLRFNSIQKLKSNQFAGLNQLIWLHNYISSVDEAFQGIRRLKELILSSNK
ITYLHNKTФHPVPNLRNLDLSYNKLQTLQSEQFKGLRKLIILHRSNSLKVPIRVFQDC
RNLDFLDLGYNRLRSLSRNAFAGLLKLKEHLHLEHNQFSKINFQAHFPRLFNRLRSIYLQWNR
IRSISQGLTWTWSSLHNLDLSGNDIQGIEPGTFKCLPNLQKLNLDNSNKLTNISQETVNAW
ISLISITLSGNMWECSRСICPLFYWLKNFKGNKESTMICAGPKHIQGEKVSDAVETYNIC
SEVQVVNTERSHLVPQTPQKPLIIPRPTIFKPDVTQSTFETPSPSPGFQIPGAEQEYEHV
SFHKIIAGSVALFLSVAMILLVIYVSWKRYPASMQLQQHSLMRRRKARESERQMNSP
LQEYYVVDYKPTNSETMDISVNGSGPCTYTISGSRECEMPHHMKPLPYYSYDQPVIGYCQA
HQPLHVTKGYETVSPEQDESPGLELGRDHSFIATIARSAAPAIYLERIAN
```

Important features of the protein:

Signal peptide:

Amino acids 1-30

Transmembrane domain:

Amino acids 425-443

N-glycosylation sites:

Amino acids 58-62;126-130;291-295;501-505

Tyrosine kinase phosphorylation site:

Amino acids 136-143

N-myristoylation sites:

Amino acids 29-35;61-67;247-253;267-273;271-277;331-337;
502-508;512-518;562-568

Glycosyl hydrolases family:

Amino acids 310-319

73/246

FIGURE 71

TTCCAGTCAGAGTTAAGTTAAAACAGAAAAAGGAAG**ATGG**CAAGAATATTGTTACTTTCC
TCCCGGGTCTTGTGGCTGTATGTGCTGTGCATGGAATATTATGGACCGTCTAGCTTCCAAG
AAGCTCTGTGCAGATGATGAGTGTGTCTATACTATTCTCTGGCTAGTGCTCAAGAAGATTA
TAATGCCCGGACTGTAGATTCAACGTTAAAAAAGGGCAGCAGATCTATGTGTACTCAA
AGCTGGTAAAAGAAAATGGAGCTGGAGAATTTGGGCTGGCAGTGTATGGTGTACCA
GACGAGATGGGAGTCGTGGTTATTCACCAGGAACCTGGTCAAGGAACAGCGTGTACCA
GGAAGCTACCAAGGAAGTTCCCACCGGATATTGACTTCTGCGAG**TAATAAATTAGTT**
AAA**ACTGCAAATAGAAAGAAAACACCAAAAATAAGAAAAGAGCAAAAGTGGCCAAAAAATG**
CATGTCTGTAATTTGGACTGACGT

74/246

FIGURE 72

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA143076
><subunit 1 of 1, 128 aa, 1 stop
><MW: 14332, pI: 4.83, NX(S/T): 0
MARILLFLPGLVAVCAVHGIFMDRLASKKLCADDECVYTISLASAQEDYNAPDCRFINV
KKGQQIYVYSKLVKENGAGEFWAGSVYGDGQDEMGVVGYFPRNLVKEQRVYQEATKEVPT
TDIDFFCE

Important features of the protein:

Signal peptide:

Amino acids 1-14

N-myristoylation site:

Amino acids 84-90

75/246

FIGURE 73

CTCAGATTTGCCATGGAGAAATTTCAAGTCGGCAATCCTGCTTCTGTGGCCATCTCTGG
TACTCTGGCCAAGACACCAAGTCAAATCTGGATCCAAAAGGACCCAAAGGACTCTCGAC
CCAAACTACCCCAGACCCCTGTCCAGAGGTTGGGGAGATCAGCTCATCTGGACTCAGACTTAC
GAAGAAGCCTTATACAAATCCAAGACAAGCAACAGACCCCTGATGGTCATTCATCACTTCCA
CGAATGCCCGCACAGTCAAGCTTAAAGAAAGTGTGCTGAAATAAGGAGATCCAGAAATTG
GCAGAGCAGTTGTTCTCCTCAACTTGATCTATGAAACAACGTGACAAGCACCTTCTCCTGA
TGGCCAGTACGTCCCCAGAATTGTGTTGTGGACCCCTCCCTGACGGTGAGGGCAGACATCA
CCGGAAGATACTCAAACCGTCTACGCTTATGAACCTCTGACACAGCTCTGTTGCACGAC
AACATGAAGAAAGCTCTCAAGTTGCTGAAGACAGAGTTGTAGAGTCAACTGTACAGTGCCTC
AGGAGCCGGGAAGGCAGAAGCAGTGTGGACCTGCCGATGACATTACAGTTAATGTTACAAC
AAATGTATTTTAAACACCCACGTGTGGGGAAACAATATTATTACTACAGACACATG
ATTTCTAGAAAATAAGTCTTGTGAGAACTCCAAA

76/246

FIGURE 74

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA143294
><subunit 1 of 1, 175 aa, 1 stop, 1 unknown
><MW: 19888.97, pI: 9.08, NX(S/T): 0
MEKFSVSAILLVAISGTLAKDTTVKSGSKKDPKDSRPKLPQTLSRGWDQLIWTQTYEE
ALYKSKTNSRPLMVIHHLDECPHSQALKKVFAENKEIQKLAEQFVLLNLIYETTDKHLSP
DGQYVPRIVFVDP SLTVRADITGRYSNRLYAYEPSDTALLHDNMKKALKLLKTEL
```

Important features of the protein:

Signal peptide:

Amino acids 1-20

77/246

FIGURE 75

GGCGGCGCCAGGGCAGGCAGGGCGGCTGGCAGCTGTGGCGCCGAC**ATGGCTGCGCTGGTGGAG**
CCGCTGGGGCTGGAGCGGGACGTGTCCGGCGGTTGAGCTCCTCGAGCGGCTCCAGCGCAG
CGGGGAGCTGCCGCCAGAACGCTGCAGGCCCTCCAGCGAGTTCTGCAGAGGCCCTCTGCT
CCGCTATCCGAGAGGTGTATGAGCAGCTTATGACACGCTGGACATCACCGGCAGCGCCGAG
ATCCGAGCCATGCCACAGCCAAGGCCACAGTGGCTGCCTCACAGCCAGCGAGGGCCACGC
ACATCCCAGGGTAGTGGAGCTACCCAAAGACGGATGAGGGCTAGGCTCAACATCATGGGTG
GCAAAGAGCAAAACTCGCCCATCTACATCTCCGGGTATCCAGGGGTGTGGCTGACCGC
CATGGAGGCCTCAAGCGTGGGATCAACTGTTGTCGGTGAACGGTGTGAGCGTTGAGGGTGA
GCAGCATGAGAAGGCGGTGGAGCTGCTGAAGGCAGGCCAGGGCTCGGTGAAGCTGGTTGTCC
GTTACACACCGCGAGTGCTGGAGGAGATGGAGGCCGGTTCGAGAAGATGCGCTCTGCCCGC
CGCGCCAACAGCATCAGAGCTACTCGTCCTGGAGTCTCGAGGT**TGA**AAACCACAGATCTGG
ACGTTCACGTGCACTCTTCTGTACAGTATTATTGTTCTGGCACTTATTAAAGATA
TTGACCTCAAAAAAAAAAAAAAA

78/246

FIGURE 76

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA143514
><subunit 1 of 1, 207 aa, 1 stop
><MW: 22896, pI: 8.93, NX(S/T): 0
MAALVEPLGLERDVSRASVELLERLQRSGELPPQKLQALQRVLQSRFCASAIREVYEQLYDT
LDITGSAEIRAHATAKATVAAFTASEGHAHPRVVELPKTDEGLGFNIMGGKEQNSPIYIS
RVIPIPGVADRHGGLKRGDQLLSVNGVSVEGEQHEKAVELLKAAQGSVKLVVRYTPRVLEE
MEARFEKMRSAARRRQQHQSYSSLESRG

Tyrosine kinase phosphorylation site:
Amino acids 51-59

N-myristoylation sites:
Amino acids 102-108; 133-139

Cell attachment sequence:
Amino acids 136-139

PDZ domain (Also known as DHR or GLGF):
Amino acids 93-174

79/246

FIGURE 77

CTGTCAGCTGAGGATCCAGCCGAAAGAGGGAGCCAGGCACTCAGGCCACCTGAGTCTACTCAC
CTGGACAACTGGAATCTGGCACCAATTCTAAACCACTCAGCTTCTCCGAGCTCACACCCCCGG
AGATCACCTGAGGACCCGAGCCATTGATGGACTCGGACGAGACCGGGTTCGAGCACTCAGGA
CTGTGGTTCTGTGCTGGCTGGTCTGCTGGAGCCTGCCAGGCACACCCCCTCCCTGACTC
CAGTCCTCTCCTGCAATTGGGGGCCAAGTCCGGCAGCGGTACCTACACAGATGATGCC
AGCAGACAGAAGCCCACCTGGAGATCAGGGAGGATGGGACGGTGGGGGGCGCTGCTGACCAG
AGCCCCGAAAGTCTCCTGCAGCTGAAAGCCTTGAAGCCGGAGTTATTCAAATCTTGGGAGT
CAAGACATCCAGGTTCCCTGTGCCAGCGCCAGATGGGGCCCTGTATGGATCGCTCCACTTG
ACCCTGAGGCCCTGCAGCTTCCGGGAGCTGCTTCTGAGGACGGATAACAATGTTACCAAGTCC
GAAGCCCACGGCCTCCGCTGCACCTGCCAGGGAAACAAGTCCCCACACCAGGACCCCTGCACC
CCGAGGACCAGCTCGTCCCTGCCACTACCAGGCCCTGCCCTGGGGACTCCCGGAGCCACCCG
GAATCCTGGCCCCCCCAGCCCCCGATGTGGGCTCCTCGGACCCCTGTGAGCATGGTGGGACCT
TCCCAGGGCGAAGCCCCAGCTACGCTTCTGAGCCAGGGCTGTTACTATGACATCTCC
TCTTATTTATTAGGTATTATCTTATTATTTTTTACTTGAGATAATAAGA
GTTCCAGAGGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAG

80/246

FIGURE 78

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA144841
><subunit 1 of 1, 208 aa, 1 stop
><MW: 22187, pI: 5.08, NX(S/T): 1
MDSDETGFEHSGLWVSVLAGLLGACQAHPPIPDSPLLQFGGQVRQRYLYTDDAQQTAEHL
EIREDTGVTGGAADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPEAC
SFRELLLEDGYNVYQSEAHGLPLHLPGNKSPHRDPAPRGPARFLPLPGLPPALPEPPGIL
APQPPDVGSQDPLSMVGPSQGRSPSYAS
```

Important features of the protein:

Signal peptide:

Amino acids 1-27

N-myristoylation sites:

Amino acids 12-18;20-26;23-29;66-72;94-100;107-113;168-174

Prokaryotic membrane lipoprotein lipid attachment site:

Amino acids 15-26

HBGF/FGF family proteins:

Amino acids 57-73;80-131

FIGURE 79

82/246

FIGURE 80**Protein File:**

MW: 73502.97, PI: 9.26

MSSQPAGNQTSPGATEDSYGSWYIDEPQGGEELQPEGEVPSCHTSIPPGLYHACLASLS
ILVLLLLAMLVRRRQLWPDCVRGRPGLPSPVDFLAGDRPRAVPAAVFMVLLSSLCLLPD
EDALPFLTLASAPSQDGKTEAPRGAWKILGLFYYAALYYPLAACATAGHTAAHLLGSTLS
WAHLGVQVWQRAECPQVPKIKYYSLLASLPLLLGLGFLSLWYPVQLVRSFSRRTGAGSK
GLQSSYSEEYLRNLLCRKKLGSSYHTSKHGFLSWARVCLRHCITYTPQPGFHLPLKLVLSA
TLTGTAIYQVALLLLGVVVPTIQKVRAGVTTDVSYLLAGFGIVLSEDKQEVELVKHHLW
ALEVCYISALVLSCLLTFLVLMRSLVTHRTNLRALHRAALDLSPHRSPHPSRQAIFCW
MSFSAYQTAFICLGLLQQIIFFLGTTALAFLVLMPLHGRNLLLFRSLESSWPFWLTIA
LAVILQNMAAHWVFLETHDGHPQLTNRRVLYAATFLLFPLNVLGAMVATWRVILSALYN
AIHLGQMDLSLLPPRAATLDPGYYTYRNFLKIEVSQSHPAMTAFCSLLLQAQSLLPRTMA
APQDSLRLPGEEDEGMQLLQTKDSMAKGARPGASRGRARWGLAYTLLHNPTLQVFRKTALL
GANGAQP

Important features of the protein:**Transmembrane domains:**

Amino acids 54-69;102-119;148-166;207-222;301-320;
364-380;431-451;474-489;512-531

N-glycosylation site:

Amino acids 8-12

N-myristoylation sites:

Amino acids 50-56;176-182;241-247;317-323;341-347;525-531;
627-633;631-637;640-646;661-667

Prokaryotic membrane lipoprotein lipid attachment site:

Amino acids 364-375

ATP/GTP-binding site motif A (P-loop):

Amino acids 132-140

FIGURE 81

AAAAAAATACAGCAGGTGAAGGAGGTTGGAGAGTAGGGGGTGGAGGGCCCACGCAGCACTTGT
CCTTCACCCCTGGAGGGGATCTGTTACATGCCAGATTGCTGGTCCCCTAGAAATGTTACTG
AGGCAGCCTCTGCATTTGCAGGGATTGTTCTACTGTTGACATTACGTAACCTCCTA
ACGCTGTCTGGGAAGATGCTACCCCTGCTCTCCCGTCTTCCTGCACCTCAGCAATGG
GATGGGCTGACTGATGCCCTGTTGGCTGGAAAGCTGACCACAGTTGCTGCAGACCAGACCCC
CTCACATAGTGAGTGCTGGCTGAGGAATCCAGGGAGAGCCCCGAGGGGGGACACTGAAGGTGT
ATCGTTGGCCCTGCCAGCTGCAAGTGAACTGTTCTGATGAATTAAAGGGAGAAAGAAG
TATTGCTAAGAATGGCAATCCTGACGCTCAGCCTCAACTCATCTTGTATTAAATACCATC
AATATCCCAGGAGCTCATAAAACGAGTCTTCTTGGAAACATGACCAAGATTGGCAA
ACGTCTCCAACATGACTTCAGCAACGGAAAACTAAGAGTCAAAGGCATTATTACCGGAAT
GCCGACATTGCTCTCGACATCGCGTAACCTCAGCAGGCCTAACTCTGCAGGACCTTCAGCT
ATGGTGTAAATTGAGGTCAGTGGCCAGAGGACAGATCCCGTCTACATTTGAGTGAAGCGGAGA
GCTACTGCAGGGTTCTGAGCAGACTCTAAATTATTTAGAAGAATCATCATGGCTCCTA
GATTAGGAATAAAACGAAGGGCCCAGGGATGGAAACGATGAGTCCAGTTGGTTACTGCAA
AGATCCAGGCCAGAAATCCAGGCACAGTGGCACACACCTGAGTCCAGATAATTCCACCTAC
TGGTCTGCTCTGTCCTACTGGTCCGAGTCCAGCCCCGACTGATTCTGGCCTGTAATG
TCTAAAAACGCTCCCTGCTGATGTTTGCAAGTGACTGTGTTACTTGAAGGCAGTTCCCTAGG
ATAAAACTAGTCGCTTATCATTACAGAATCATTCACTGAGCATCAACTATGTAACCAGCATT
GGGTTGGGTGCCAGAGATCCAAAGCTAACGACACCAAAACTGCTCTCCAGGAAACGAGAGGC
TGAGAA

84/246

FIGURE 82

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA149995
><subunit 1 of 1, 95 aa, 1 stop
><MW: 10704, pI: 10.00, NX(S/T): 2
MAILTLSIQLILLIPIPSISHEAHKTSLSWKHDQDWANVSNMTFSNGKLRVKGIYYRNAD
ICSRHRVTSAGLTLQDLQLWCNLRSVARGQIPSTL
```

Important features of the protein:

Signal peptide:

Amino acids 1-19

N-glycosylation sites:

Amino acids 38-42; 41-45

N-myristoylation site:

Amino acids 89-95

FIGURE 83

AATAGAAGTCCTCAGGACGGAGCAGAGGTGGCGGGCCGGCTGACTGCGCTCTGTT
TCTTCCATAACCTTCTCGACTCGAACATCAGGCTGCTCGAAGGGCTAGTCCGGA
CACTAGGGTGCCGAACCGCTGATGCCCGAGTGCCTGCAGGGCTTCCGCTAACCAATGCT
GCCGCCGCCGCCGCCAGCTGCCCTGGCGCTGCCCTGCTCCTGCTACTGCTGGTGGTGC
TGACGCCGCCCGACCAGCGCAAGGCCATCCCCAGGGCCAGATTACCTGCAGGCGCCAG
ATGCGGCTGCTAGCGGAGGGCGAGGGCTGCCTCCCTGCCAGAAGAGTGCAGGCCGCC
GCCGGGCTGCCCTGGCGGCCAGGGTGCAGCGTGCAGGCTGCTGGAAATGCCAAC
TCGAGGGCAGCTCTGCACCTGGACCCCCAGTGCTCACTTCTACGGGACTGCGGCGAGCAG
CTTGAGTGCCGGCTGGACACAGGGCGACCTGAGCCGAGAGGGTGCAGGAAACCTCTGTG
TGCCTGTCGTTCGCAGAGTCCGCTCTGCAGGGTCCGACGGTCACACCTACTCCCAGATCTGCC
GCCCTGCAGGAGGCCGCCGCTGGCCCGATGCCAACCTCACTGTCAGGACACCCGGGCCC
TGCAGATCGGGCCCCAGATCGTGTACATCCATATGACACTTGGAAATGTGACAGGGCAGGA
TGTGATCTTGGCTGTGAAGTGTTCGCCTACCCATGCCCTCCATCGAGTGGAGGAAGGATG
GCTTGGACATCCAGCTGCCAGGGATGACCCCCACATCTGTGCAGTTAGGGTGGACCC
CAGAGGTTGAGGTGACTGGCTGGCTGCAGATCCAGGCTGTGCGTCCAGTGATGAGGGCAC
TTACCGCTGCCCTGGCGCAATGCCCTGGCTCAAGTGGAGGCCCTGCTAGCTGACAGTGC
TCACACCTGACCAGCTGAACCTACAGGCATCCCCAGCTGCATCAAAACCTGGTTCCT
GAGGAGGAGGCTGAGAGTGAAGAGAATGACGATTACTACTAGGTCCAGAGCTGGCCATG
GGGGTGGGTGAGCGGCTATAGTGTTCATCCCTGCTCTGAAAAGACCTGGAAAGGGGAGCAG
GGTCCCTTCATCGACTGCTTCATGCTGTCAGTAGGGATGATCATGGAGGCCTATTGACT
CCAAGGTAGCAGTGTGGTAGGATAGAGACAAAGCTGGAGGAGGGTAGGGAGAGAAGCTGAG
ACCAGGACCGGTTGGGTACAAGGGCCCATGCAGGAGATGCCCTGGCCAGTAGGACCTCCA
ACAGGGTTTCCCAGGCTGGGTGGGGCCTGAGCAGACACAGAGGTGCAGGCACCCAGGAT
TCTCCACTTCTCCAGCCCTGCTGGGCCACAGTTCAACTGCCCTCCTCCCAGGCCCTGGT
TCTTGCTATTCTGGTCCCCAACGTTATCTAGCTTGTGGCCCTTCCCCAAACTCATCT
TCCAGAACTTTCCCTCTCTAACGCCCCAGTGCACCTACTAACGACTGAGTCCCTTTGCT
GTCTGCCGTCTTGTACAAGAGAGAACAGCGGAGCATGACTTAGTTAGTCAGTGCAGAGAGA
TTT

86/246

FIGURE 84

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA167678
><subunit 1 of 1, 304 aa, 1 stop
><MW: 32945, pI: 4.69, NX(S/T): 3
MLPPPRPAAALALPVLLLLLVLTPPPTGARPSPGPDYLRRGWMRLLAEEGEGCAPCRPEE
CAAPRGCLAGRVRDAGCCWECANLEGQLCDLDPASAHFYGHCGEQLECRLDTGGLLSRGE
VPEPLCACRSQSPLCGSDGHTYSQICRQLQEAARARP DANLTVAHPGPGCESGPQIVSHPYD
TWNVTGQDVIFGCEVFAYPMASIEWRKDG LDIQLPGDDPHISVQFRGGPQRFEVTGWLQI
QAVRPSDEGTYRCLGRNALGQVEAPASLTVDQLNSTGIPQLRSLNLVPEEEAESEEN
DDYY
```

Important features of the protein:**Signal peptide:**

Amino acids 1-30

N-glycosylation sites:

Amino acids 159-163;183-187;277-281

Tyrosine kinase phosphorylation site:

Amino acids 244-252

N-myristoylation sites:

Amino acids 52-58;66-72;113-119;249-255

Kazal-type serine protease inhibitor domain:

Amino acids 121-168

Immunoglobulin domain:

Amino acids 186-255

Insulin-like growth factor binding proteins:

Amino acids 53-90

87/246

FIGURE 85

CAAAGCGGCGGCTGTCCCGGGTGCCTGGCTGGGGCGGAGAGGCGGCGTGGCTCCCTGGGG
TGTGTGAGCCCGGT**TATGG**AGCCGGGCCACAGCCGCGAGCGGAGGGTGTTCGTTGCCGCC
GTGGCTGCCGCTGGGCTGCTGTGGTCGGGCTGGCCCTGGCGCCTCCCTGGCA
GCAGTCGCACAGGGCTTCCACGACCTCCGTGGAGCAGCAGTGTGGAGGGTGGAGGAC
TTGTCCTGTCCTCCTGCAGGGTGGAGGGCTGGGCTGTGCGTGCCTGGGACCTGCC
GGATCTGGATCCTGAGTGCAGGGAGCTCTGCTGGACTTCGCAACAGCAGCGAGAGCTGA
CAGGGTGTCTGGTGCAGCAGCGCCGGCCGTGCCTCTGTCAGACCTGCTACCCCTCTC
CAACAGGTCGTCAGCAAGATGGACAACATCAGCCGAGCCGCGGGGAATACTCAGAGAGTCAG
AGTTGTGCCAGAAGTCTCTTAATGGCAGATAGAATGCAAATAGTTGTGATTCTCAGAATT
TTTAATACCACATGGCAGGAGGCAAATTGTGCAAATTGTTAACAAACAGTGAAGAAT
TATCAAACAGCACAGTATATTCTTAATCTATTAAATCACACCCCTGACCTGCTTGAAACAT
AACCTTCAGGGGAATGCACATAGTCTTTACAGACAAAAAATTATTCAAAGTATGCAAAA
CTGCCGTGAAGCATAACAAACTCTGAGTAGTCTGTACAGTGAATGCAAAAATGAATGAAC
TTGAGAATAAGGCTGAACACATTTATGCATTGATGTGGAAGATGCAATGAACATC
ACTCGAAAATATGGAGTCGAACTTCAACTGTCAGTCCCTGCAGTGACACAGTGCCTGT
AATTGCTGTTCTGTGTTCACTCTTCTACCTGTTGTCTACCTTAGTAGCTTCTTC
ACTCAGAGCAAAAGAACGCAAACCTCATCTGCCAAACGTCTCAAGTCCAGTACAGTTT
GCAAATATTCAAAGGAAATTCAAAC**TGA**AGACCTACAAATGGAGAATTGACATATCAGTGAA
TGAATGGTGGAAAGACACAACCTGGTTTCAGAAAGAAGATAAACTGTGATTTGACAAGTCAAG
CTCTTAAGAAATACAAGGACTTCAGATCCATTAAATAAGAATTTCGATTTCTTCC
TTTCCACTCTTCTAACAGATTGGATATTAAATTCCAG

88/246

FIGURE 86

```
>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA168028
><subunit 1 of 1, 334 aa, 1 stop
><MW: 37257, pI: 5.95, NX(S/T): 10
MEPGPTAAQRRCSLPPWLPLGLLLWSGLALGALPFGSSPHRVFHDLLSEQQLLEVEDLSL
SLLQGGGLGPLSLPPDLPDLDPECRELLDFANSSAELTGCLVRSARPVRLCQTCYPLFQ
QVVKMDNISRAGNTSESQSCARSLLMADRMQIVVILSEFFNTTWQEANCANCLNNSE
ELSNSTVYFLNLFNHTLTCFEHNLQGNAHSLLQTKNYSEVCKNCREAYKTLSSLYSEMQK
MNELENKAEPGTHLCIDVEDAMNITRKLWSRTFNCSVPCSDTVPVIAVSVFILFLPVVFY
LSSFLHSEQKKRKLILPKRLKSSTSFAPIQENSN
```

Important features of the protein:**Signal peptide:**

Amino acids 1-31

Transmembrane domain:

Amino acids 278-300

N-glycosylation sites:

Amino acids 93-97;128-132;135-139;163-167;177-181;
184-188;194-198;216-220;263-267;274-278

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 10-14

N-myristoylation sites:

Amino acids 27-33;206-212;251-257

Leucine zipper pattern:

Amino acids 190-212

FIGURE 87

**ATGCTGGTAGCCGGCTTCCCTGCTGGCGCTGCCGAGCTGGGCCGCGGGCAGCCCCCAGGGC
GGCAGGCGCCCCCGCGCGCCGCGGGCTGGCGGGCTGCTCAGTGCCTTACACAGCTGAGCTGGAGCAGC
TGTACGGGCGCTGGCGGCCGGCGTGTCTCAGTGCCTTACACAGCTGAGCTGGGGCCG
CGTGAAGCAGGCGCAACCGAGCTGCCGGCAGGGGCAGGCCGGCACCGCCGTTCCG
GCCGCCACCAACCTGCGCAGCGTGTGCCCTGGGCCTACAGAAATCTCCTACGACCAGGCGA
GGTACCCCAGGTACCTGCTGAAGCCTACTGCCGTGCCGGGCTGCCTGACCAGGCTGTT
GGCAGGAGGACGTGCGCTTCCGCAAGCAGGCCCTGTCTACATGCCAACGTCGCTTGC
CACCCCCGCTGCGCCGGGCCGTTCCGTACACCGAGGCCTACGTCAACATCCCCGTGG
GCTGCACCTGCGTCCCCGAGCGGAGAAGGACGCAGACAGCATCAACTCCAGCATCGACAAA
CAGGGCGCCAAGCTCCTGCTGGGCCCAACGACGCGCCGCTGGCCCCCTGAGGCCGTC
CCCCGGAGGTCTCCCCGGCCGCATCCCGAGGCGCCAAGCTGGAGGCCCTGGAGGGCTC
GGTGGCGACCTCTGAAGAGAGTGCACCGAGCAAACCAAGTGCCGGAGCACAGCGCCGCT
TTCCATGGAGACTCGTAAGCAGCTTCATCTGACACGGCATCCCTGGCTTGTCTTAACTAC
AAGCAAGCAGCGTGGCTGGAAGCTGATGGGAAACGACCCGGCACGGCATCCTGTGCGGC
CCGCATGGAGGGTTGGAAAAGTTACGGAGGCTCCCTGAGGAGCCTCTCAGATCGCTGCT
GCGGGTGCAGGGCGTGAACCTACCGCTGGGTGCTTGCCAAAGAGATAGGGACGCATATGCTTT
TTAAAGCAATCTAAAAATAATAATAAGTATAGCGACTATATACCTACTTTAAATCAACTG
TTTGAATAGAGGGCAGAGCTATTTATATTCAAATGAGAGCTACTCTGTTACATTCTTA
ACATATAAACATCGTTTTACTCTCTGGTAGAATTAAAGCATAATTGGAAATCCTT
GGATAAAATTGAGCTGGTACACTCTGGCCTGGGTCTCTGAATTAGCCTGTCACCGATGG
CTGACTGATGAAATGGACACGCTCATCTGACCCACTCTCCTCCACTGAAGGTCTCAGC
GGCCTCCAGGTGGACCAAGGGATGCACAGGGGGCTCCATGCCCAAGGGCCAGCTAAGAGT
TCCAAAGATCTCAGATTGGTTTAGTCATGAATACATAAACAGTCTCAAACCTGCACAATT
TTTCCCCCTTGAAAGCCACTGGGGCCAATTGTGGTTAAGAGGTGGTGAGATAAGAAGT
GGAACGTGACATCTTGCCAGTTGTCAAGAAGATCCAAGCAGGTATTGGCTTAGTTGAAGG
GCTTCTGGATCAGGCTGAATATGAGGACAAAGTGGGCCACGTTAGCATCTGCAGAGATCAAT
CTGGAGGCTCTGTTCTGCATTCTGCCACGAGAGCTAGGTCTGATCTTCTTAAAGT
GAAAGTCTGTCTGAACACAATTATTGTAAAAGTTAGTAGTTCTTTAAATCATTAAA
AGAGGCTGCTGAAGGGAT**

90/246

FIGURE 88

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA173894
><subunit 1 of 1, 202 aa, 1 stop
><MW: 21879, pI: 9.30, NX(S/T): 2
MLVAGFLLALPPSWAAGAPRAGRRPARPRGCADRPEELLEQLYGRLLAAGVLSAFHHTLQL
GPREQARNASCPAGGRPGDRRFRPPTNLRSVSPWAYRISYDPARYPRYLPEAYCLCRGCL
TGLFGEEDVRFRSAPVYMPVVLRRTPACAGGRSVYTEAYVTIPVGCTCVPPEKDADSI
NSSIDKQGAKLLGPNDAPAGP
```

Important features of the protein:

Signal peptide:

Amino acids 1-15

N-glycosylation sites:

Amino acids 68-72;181-185

Tyrosine kinase phosphorylation site:

Amino acids 97-106

N-myristoylation sites:

Amino acids 17-23;49-55;74-80;118-124

Amidation site:

Amino acids 21-25

FIGURE 89

CCGGGGCCTCCGGAGAACGCTGTCCCAGAACGTGCAGGGAGCGGCCCGCGTCCGCGCG
TCCCCCGCTCCCTGGCAATTCCCGACTTCCCAACGGCTTCCCGCTGGCAGCCCCGAAGCCGC
ACCATGTCCGCCTCGGTTGCTGCTGGCGGGCTCTGCAGGGCTCTGGCGTCAAGACCCGGT
TTTCAAAATTCACTTCTACAGATCGTAATTCCAGAGAAAATCAAACAAATACAAATGACAG
TTCAGAAATAGAATATGAACAAATATCCTATATTATTCAATAGATGAGAAACTGTACACTG
TGCACCTAAACAAAGATATTAGCAGATAATTATGATCTATTGTACAATCAAGGA
TCTATGAATACTTATTCTCAGATATTCACTCAATGCTACTATCAAGGAATATTGAAGG
ATATCCAGATTCCATGGTCACACTCAGCACGTGCTGGACTAAGAGGAATACTGCAATTG
AAAATGTTCTTATGAAATTGAGCCTCTGGAACTGCAAGTTGAATTTCAGCATGTTCTTAC
AAATTAAAGAATGAAGACAATGATATTGCAATTATTGACAGAAGCTGAAAGAACAA
AATGGATGACAACATTTATAAGTAAAAATCAGAACAGCTGTTCCAGATTATTCCCTC
TTTATCTAGAAATGCATATTGTTGGGAGACAAACTTGTATGATTACTGGGCTCTGATAGC
ATGATAGTAACAAATAAGTCATCGAAATTGTTGGCCTTGCAATTCAATGTTACCCAAATT
TAAAGTTACTATTGTGCTGTCACTATTGGAGTTATGGTCAGATGAAAATAAGATTCTACAG
TTGGTGGGCAGATGAATTATTGCAAAATTGAGATGAAACAACTTATCTTAC
AGGCCTCATGATATTGCATATCTACTAATTATGGATTATCCTGTTATTGGGAGCAGT
GTTCTGGAACAATGTGTATTACTCGTTACTGCAGGAGTTGCATTGTACCCAAAGGAGA
TAACCTGGAGGCATTGCAGTTATTGTCACCCAGATGCTGGCACTCAGTCTGGAAATATCA
TATGACGACCCAAAGAAATGTCAATGTTCAAGTCCACCTGTATAATGAATCCAGAAAGTTGT
GCAATCCAATGGTGTGAAGACTTTAGCAGTTGCAGTTGAGGAGCTTCAAAATTCAATT
CAAATGTGGGTGTCAAATGTCTCAGAATAAGCCACAAATGCAAAAAAAATCTCGAAACCA
GTCTGTGGCAATGGCAGATTGGAGGGAAATGAAATCTGTGATTGTGGTACTGAGGCTCAATG
TGGACCTGCAAGCTGTTGATTTCGAACTGTGTACTGAAAGACGGAGCAAATGTTATA
AAGGACTGTGCTGCAAAGACTGTCAAATTACATCAGGCGTTGAATGTAGGCCGAAAGCA
CATCCTGAATGTGACATCGCTAAAATTGTAATGAAAGCTCACCAGAATGTGGCCTGACAT
AACTTTAATCAATGGACTTCATGCAAAATAATAAGTTATTGTTATGACGGAGACTGCC
ATGATCTCGATGCACGTTGAGAGTGTATTGGAAAAGGTTCAAGAAATGCTCCATTGCC
TGCTATGAAGAAATACAATCTCAATCAGACAGATTGGAACACTGTGGTAGGGATAGAAATAA
CAAATATGTGTTCTGTTGAGGAAATCTTATATGTGAAAGATTAGTTGTACCTACCCCTA
CTCGAAAGCCTTCCATCAAGAAAATGGTGTGATTATGCTTGTACGAGATTCTGTA
TGCATAACTGTAGACTACAAATTGCCTCGAACAGTCCAGATCCACTGGCTGTCAAAATGG
CTCTCAGTGTGATATTGGGAGGGTTGTGTAATCGTGAATGTGTAGAATCAAGGATAATTAG
GCTTCAGCACATGTTGTTCAACACAGTGTCTGGACATGGAGTGTGATTCCAGAAACAA
GTGCCATTGTTGCCAGGCTATAAGCCTCAAACACTGCCAAATACGTTCAAAGGATTTC
TATTTCCTGAGGAAGATAAGGTTCAATCATGAAAGAGCATCTGGGAAAGACTGAAAACACC
TGGCTTCTAGGTTCTCATTGCTCTTCTATTCTCATTGTAACAACCGCAATAGTTGGC
AAGGAAACAGTTGAAAAAGTGGTGTGCAAGGAAGAGGAATTCCAAGTAGCGAATCTAAAT
CGGAAGGTAGCACACAGACATATGCCAGCCAATCCAGCTCAGAAGGCAGCACTCAGACATAT
GCCAGCCAACCAAGATCAGAAAGCAGCAGTCAGCTGATACTAGCAATCCAATCAGAAGA
TAGTGCTGAAGCATATACTAGCAGATCCAATCACAGGACAGTACCCAAACACAAAGCAGTA
GTAACTAGTGATTCCCTCAGAAGGCAACGGATAACATCGAGAGTCTCGCTAAGAAATGAAA
TTCTGTCTTCCCTCGTGGTCACAGCTGAAAGAAACAATAATTGAGTGTGGATC

92/246

FIGURE 90

>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA176775
><subunit 1 of 1, 787 aa, 1 stop
><MW: 87934, pI: 5.49, NX(S/T): 4
MFRLWLLLAGLCGLLASRPGFQNSLLQIVIPEKIQTNTNDSSEIEYEQISYIIPIDEKLY
TVHLKQRYFLADNFMIFYLYNQGSMNTYSSDIQTQCYQGNIEGYPDSMVTLSTCSGLRGI
LQFENVSYGIEPLESAVEFQHVLYKLKNEDNDIAIFIDRSLKEQPMDDNIFISEKSEPAV
PDLFPLYLEMHIVVDKTLYDYWGSDSMIVTNKIEIVGLANSMFTQFKVTIVLSSLELWS
DENKISTVGEADELLQKFLEWKQSYLNLRPHDIAYLLIYMDYPRYLGAVFPGTMCITRYS
AGVALYPKEITLEAFAVIVTQMLALSLGISYDDPKKCQCSESTCIMNPEVVQSNGVKTF
SCSLRSFQNFISNVGVKCLQNKPQMOKKSPKPVCGNGRLEGNEICDCGTEAQCGPASCD
FRTCVLKDGAKCYKGLCCKDCQILQSGVECRPKAHPECEDIAENCNGSSPECGPDTILING
LSCKNNKFICYDGDCHDLDARCESVFGKGSRNAPFACYEEIQSQSDRFGNCGRDRNNKYV
FCGWRNLICGRLVCTYPTRKPFHQENGDIYAFVRDSVCITVDYKLPRTPDPLAVKNGS
QCDIGRVCVNRECVERIICKASAHVCSQQCSGHGVCDSRNKCHCSPGYKPPNCQIRSKGF
SIFPEEDMGSIMERASGKTENTWLLGFLIALPILIVTTAIVLARKQLKKWFAKEEEFPSS
ESKSEGSTQTYASQSSSEGSTQTYASQTRSESSSQADTSKSKESEDSEAAYTSRSKSQDST
QTQSSSN

Important features of the protein:**Signal peptide:**

Amino acids 1-16

Transmembrane domain:

Amino acids 309-326; 681-705

N-glycosylation sites:

Amino acids 39-43; 125-129; 465-469; 598-602

Glycosaminoglycan attachment site:

Amino acids 631-635

Tyrosine kinase phosphorylation site:

Amino acids 269-276

N-myristoylation sites:Amino acids 13-19; 82-88; 99-105; 218-224; 401-407; 634-640;
726-732; 739-745**EGF-like domain proteins:**

Amino acids 642-654

Disintegrins proteins:

Amino acids 400-407; 422-472; 403-453; 467-517; 634-684

Reprolysin (M12B) family zinc metalloprotease:

Amino acids 186-383

Reprolysin family propeptide:

Amino acids 63-176

FIGURE 91

CAACAGACAGCACTCCAGCACTCTGTTGGGGGGATTGAAACAGCAAAATCACTCATAAA
AGCAAAAAATTGCAAAAAAAATAGTAATAACCAGCATGGCACTAAATAGACCATGAAAAG
ACATGTGTGCAGTATGAAATTGAGACAGGAAGGCAGAGTGTCAAGCTTGTCCACCTCAG
CTGGGAATGTGCATCAGGCAACTCAAGTTTCACCACGGCATGTGTCTGTGAATGTCCGCA
AAACATTCTCTCTCCCCAGCCTCATGTGTTAACCTGGGATGATGTGGACCTGGCACTGTGG
ATGCTCCCTCACTCTGCAAATTCAAGCTGGCAGCTCTGCCAGCTAAGCCTGAGAACATTTC
CTGTGTCTACTACTATAGGAAAAATTAAACCTGCACTTGGAGTCCAGGAAAGGAAACAGTT
ATACCCAGTACACAGTTAAGAGAACTTACGCTTTGGAGAAAACATGATAATTGTACAAACC
AATAGTTCTACAAGTAAAATCGTGCCTCGTGTCTTTTCCCTCCAAGAATAACGATCCC
AGATAATTATACCATGAGGTGGAAGCTGAAATGGAGATGGTGTAAATTAAATCTCATATGA
CATACTGGAGATTAGAGAACATAGCGAAAACCTGAACACCCTAAGATTTCGTTGAAACCA
GTTTGGGCATCAAACGAATGATTCAAATTGAATGGATAAAGCCTGAGTTGGCCCTGTTTC
ATCTGATTTAAAATACACACTTCGATTCAAGCAGTCAACAGTACCAAGCTGGATGGAAGTCA
ACTTCGCTAAGAACCGTAAGGATAAAACCAACGTACAACCTCACGGGCTGCAGCCTTT
ACAGAATATGTCATAGCTCTGCGATGTGGCTCAAGGAGTCAAAGTTCTGGAGTGACTGGAG
CCAAGAAAAATGGGAATGACTGAGGAAGAAGGCCAGTGGCTTGTATGGAAGAAGGCAAGAGGA
GCCCGAGTCCTAGAGAAAACACTTGGCTACAACATATGGTACTATCCAGAAAGCAACACTAA
CCTCACAGAAAACAATGAACACTACTAACACAGCAGCTGAACGTGCATCTGGAGGCGAGAGCT
TTTGGGTGTCTATGATTCTTATAATTCTCTGGAAAGTCTCCAGTGGCCACCTGAGGATT
CCAGCTATTCAAGAAAAATCATTCAAGTGCATTGAGGTCACTGCAGGCCTGCGTTGAGGA
CCAGCTAGTGGTGAAGTGGCAAAGCTCTGCTTAGACGTGAACACTTGGATGATTGAATGGT
TTCGGATGTGGACTCAGAGCCCACCACCCCTTCTGGAAATCTGTGTCTCAGGCCACGAAC
TGGACGATCCAGCAAGATAATTAAACCTTCTGGTCTATAACATCTGTGTATCCAAT
GTTGCATGACAAAGTGGCGAGCCATATTCCATCCAGGCTTATGCCAAAGAAGGCGTTCCAT
CAGAAGGTCTGAGACCAAGGTGGAGAACATTGGCGTGAAGACGGTCACGATCACATGGAAA
GAGATTCCCAAGAGTGGAGAGAAAGGGTATCATCTGCAACTACACCCTTACCAAGCTGA
AGGTGGAAAAGGATTCTGTAAGCACGCCATAGCGAAGTGGAAAAACCCCAAGCCCCAGA
TAGATGCTATGGATAGACCTGGTAGGCATGGCTCCCCCATCTCATGTGACTTGCAACCT
GGCATGAATCACTTAGCTTAAATCTCTGAAAATGGGCCAAGAGCACCCACCTTT
GGGTTTGGGGTTAAATGAGAGTGAAGTGAACAGTACCTGAGAGGAGAGTCTGAGGAAAT
GGAAGGAGTTGTTATAATTGCTGGTTAGGCCCTGAATTGACCTCCGGAGCTCCCCGA
CCATCATTCCCAGGAATGGCGTGCCTGGCTAAAGAGTGGAGGAGAACAGACCTGTACCA
TGACTTCTACTGCCCTGCCAAATCATGCTTTGTTTCAGTCCACCTTATCTCTGACATCT
TAAATACTGGCAAGGCTGGATTCTGCTTAGGCTAAATAATTCTTATGGTAAAATA
CACGTAAAATATTCCAGTTAAACATTGAAAGTGTACAATTAGTGGCATTAGAAGCA
TTCACAATATTGTGCAACCACCACTATTCCAGAACTCTTCTATTCTGCCAAATAGA
AGCCCTATAACCCATTAGTCACCCCCATTCTCCACAGCCCCCTGGCAACTAC
CAAACGTGTTGTCTATGGATTGCCATTGGATATTCTATACATAGAATCATAA
ANTAAAAAA

94/246

FIGURE 92

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA177313
><subunit 1 of 1, 582 aa, 1 stop
><MW: 66605, pI: 8.14, NX(S/T): 15
MCIRQLKFFTTACVCECPQNIILSPQPSCVNLGMMWTALWMLPSLCKFSLAALPAK PENI
SCVYYYRKNLCTWSPGKETSYTQYTVKRTYAFGEKHDNCTNSSTSENRASCFFLPRI
TIPDNYTIEVAENGDGVIKSHMTYWRLENIAKTEPPKIFRVKPVLGIKRMIQIEWIKPE
LAPVSSDLKYTLRFRTVNSTSWMEVNFAKNRKDKNQTYNLTGLQPFTEYVIALRCAVKE
KFWSDWSQEKMGMTEEEAPCGLELWRVVLKPAEADGRRPVRLWKKARGAPVLEKTLGYN
WYYPESNTNLTEMNTTNQQLELHLGGESFWVSMISYNSLGKSPVATLRIPIQEKSFQC
IEVMQACVAEDQLVVVKWQSSALDVNTWMIEWFPDVDSEPTTLSWESVSQATNWTIQQDKL
KPFWCYNISVYPMHLHDKVGEPEYSIQAYAKEGPSEGPETKVENIGVKTVTITWKEIPKSE
RKGIICNYTIIFYQAE GGKGFC KHAHSEVEKNPKPQIDAMDRPVVGMAPP SHCDLQPGMNH
LASNLSENGAKSTHLLGFWG LNESEVTVPERRVLRKWELL
```

Important features of the protein:**Signal peptide:**

Amino acids 1-46

N-glycosylation sites:

Amino acids 59-63;69-73;99-103;103-107;125-129;198-202;
215-219;219-223;309-313;315-319;412-416;
427-431;487-491;545-549;563-567

N-myristoylation sites:

Amino acids 32-38;137-143;483-489;550-556;561-567

Amidation site:

Amino acids 274-278

Growth factor and cytokines receptors family signature 1:

Amino acids 62-75

Fibronectin type III domain:

Amino acids 54-144;154-247

95/246

FIGURE 93

ATTCTCCTAGAGCATTTGGAAGC**ATG**AGGCCACGATGCTGCATCTGGCTTGCTGCTGCTGGATAACAGTCTCCTCCAGTGTCAAAGGAACATACAGACGCTCTGGCTCAGGACTGTGGCTGTGCCAGCGACACCCAGGTGTGGGAACAAGATCTACAACCCCTCAGAGCAGTGCTGTTATGATGATGCCATCTTATCCTAAAGGAGACCCGCGCTGTGGCTCCACCTGCACCTTCTGGCCCTGCTTGAGCTCTGCTGTCCCAGTCTTGGCCCCCAGCAGAAGTTCTTGTGAAGTTGAGGGTTCTGGGTATGAAGTCTCAGTGTCACTTATCTCCATCTCCGGAGCTGTACCAGGAACAGGAGGCACGTCTGTACCC**AT****AAA**ACCCAGGCTCCACTGGCAGACGGCAGACAAAGGGAGAAGAGACGAAGCAGCTGGACATCGGAGACTACAGTTGAACCTCGGAGAGAAGCAACTTGACTTCAGAGGGATGGCTCAATGACATAGCTTGGAGAGGAGCCCAGCTGGGATGGCAGACTTCAGGGGAAGAATGCCTCCTGCTTCATCCCCTTCCAGCTCCCCTCCGCTGAGAGCCACTTCATCGGCAATAAAATCCCCACATTACCATCT

96/246

FIGURE 94

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57700
<subunit 1 of 1, 125 aa, 1 stop
<MW: 14198, pI: 9.01, NX(S/T): 1
MPPRCCILALVCWITVFLLQCSKGTTDAPVGSGLWLCQPTPRCGNKIYNPSEQCCYDDAI
LSLKETRRCGSTCTFWPCFELCCPESFGPQQKFLVKLRLVLMKSQCHLSPISRCTRNR
HVLYP
```

Important features:

Signal peptide:

Amino acids 1-21

N-myristoylation sites:

Amino acids 33-39; 70-76

Anaphylatoxin domain proteins:

Amino acids 50-60

97/246

FIGURE 95

GCATTTTGTCTGTGCTCCCTGATCTTCAGGTACCCACCAATGAAGTTCTTAGCAGTCCTGGT
ACTCTGGGAGTTCCATTTCTGGTCTCTGCCAGAATCCGACAAACAGCTGCTCCAGCTG
ACACGTATCCAGCTACTGGTCCTGCTGATGATGAAGCCCCTGATGCTGAAACCAC TGCTGCT
GCAACCAC TGCGACC ACTGCTGCTCCTACCACTGCAACCACCGCTGCTTCTACCACTGCTCG
TAAAGACATCCAGTTACCCAAATGGGTGGGGATCTCCGAATGGTAGAGTGTGTCCCTT
GAGATGGAATCAGCTGAGTCTGCAATTGGTCACAACATTGCTGCTTCTGTGATTTC
ATCCAAC TACTTACCTGCCTACGATATCCCCTTATCTCTAATCAGTTATTTCTTCAA
ATAAAAAAATAACTATGAGCAACATAAAAAAAAAAAAAAA

98/246

FIGURE 96

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62872
<subunit 1 of 1, 90 aa, 1 stop
<MW: 9039, pI: 4.37, NX(S/T): 1
MKFLAVLVLLGVSIFLVSAQNPTTAAPADTYPATGPADDEAPDAETTAATTATTAAAPTT
ATTAASTTARKDIPVLPKWVGDLPNGRVCP
```

Important features:

Signal peptide:

Amino acids 1-19

99/246

FIGURE 97

GGACTCTGAAGGTCCAAGCAGCTGCTGAGGCCCAAGGAAGTGGTCCAACCTGGACCC
CTAGGGGTCTGGATTGCTGGTTAACAAAGATAACCTGAGGGCAGGACCCATAGGGGAATGC
TACCTCCTGCCCTTCCACCTGCCCTGGTGTTCACGGTGGCTGGTCCCTGCCGAGAGA
GTGTCCTGGGTCAAGGACGCAGAGGACGCTCACAGACTCCAGCCCTTGTACCGAGAGGAC
ACTTGGCAAGGTCCAGCGATGGTCCGGAGTCCACACACAGACTGGCAGGGCAGGGAGGGG
GACAGTTCTGTTGTGTTGGTGGACAGTAAGAGGGTCTGGCCAGTCCAGGGTGGGGCG
GCAAACCTCCATAAAGAACCAAGAGGGTCTGGGCCACAGAGTCATCTGCCAGCTCCT
CTGCTGCTGGCCAGTGGAGTGGCACGAGGTGGGCTTGTGCCAGAAAACCAAGGCTGG
ATTGCTGCGGGCCATGGTCCCTGTCTAGGGCAGCAATTCTCAACCTTCTGCTCTCAGGA
CCCCAAAGAGCTTCATTGTATCTATTGATTTTACCACTAGCAATTAAACTGAGAAAT
GGGCCGGGCACGGTGGCTCACGCCGTAAATCCCAGCACTTGGGAGGCCGAGGCCGGTGGAT
CACCTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTAAACCTTGTCTACTAAAAA
TACAAAAAAATTAGCCAGGCACAGTGGTGTGCACTGGTAGTCCAGTTACTCGGGAGGCTGAG
GCAGGAAAATCGCTTGAACCCAGGAGGCGGACGTTGCGGTGAGCCGAGATCGCGCCGCTGAT
TCCAGCCTGGCGACAAGAGTGAGACTCCATCTCACACA

100/246

FIGURE 98

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62876
<subunit 1 of 1, 120 aa, 1 stop
<MW: 12925, pI: 9.46, NX(S/T): 0
MLPPALPPALVFTVAWSLLAERVSWVRDAEDAHLQLQPFVTERTLGKVQRWSGVHTQTGGR
AGGGQFCCAWLDSKRLASPGWGAANSIKNQRVWAPATESSAQLLCCWPVGVARGGALCQ

Important features:

Signal peptide:

Amino acids 1-17

N-myristoylation sites:

Amino acids 58-64;63-69;64-70;83-89;111-117;115-121

101/246

FIGURE 99

102/246

FIGURE 100

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66660
><subunit 1 of 1, 209 aa, 1 stop
><MW: 21588, pI: 5.50, NX(S/T): 0
MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQSNQVFPSSLIPLTQML
TLGPDLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEELP
QIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSGTDD
DFAVTPAGIQRSTHAIEEATTESANGIQ
```

Important features of the protein:**Signal peptide:**

Amino acids 1-16

Leucine zipper patterns:

Amino acids 10-32;17-39

N-myristoylation sites:

Amino acids 12-18;25-31;36-42;74-80;108-114;111-117;
135-141;151-157;159-165;166-172;189-195

103/246

FIGURE 101

GGGGTCTCCCTCAGGGCCGGAGGCACAGCGGTCCCTGCTGCTGAAGGGCTGGATGTACGC
ATCCGCAGGTTCCCGCGGACTGGGGCGCCCGCTGAGCCCCGGCGCCCGCAGAAGACTTGT
GTTGCCTCCTGCAGCCTCAACCCGGAGGGCAGCGAGGGCTACCACCATGATCACTGGTGT
GTCAGCAGTCGCTTGTGGACCCCAGTGGCGTCTGACCTCGCTGGCGTACTGCCTGCACC
AGCGCGGGTGGCCCTGGCCGAGCTGCAGGAGGCCATGGCCAGTGTCCGGTGCACCGCAGC
CTGCTGAAGTTGAAAATGGTGCAGGTGCTGTTGACACCGGGCTCGGAGTCCTCTCAAGCC
GCTCCCGCTGGAGGAGCAGGTAGAGTGGAAACCCCAAGCTATTAGAGGTCCCACCCCAAACCTC
AGTTGATTACACAGTCACCAATCTAGCTGGTGGTCCGAAACCATATTCTCCTTACGACTCT
CAATACCATGAGACCACCCCTGAAGGGGGCATGTTGCTGGCAGCTGACCAAGGTGGCAT
GCAGCAAATGTTGCCTGGAGAGAGACTGAGGAAGAACTATGTGGAAGACATTCCCTTTC
TTTCACCAACCTCAACCCACAGGAGGTCTTATTGTTCCACTAACATTTTGGAAATCTG
GAGTCCACCCGTTGCTGGCTGGCTGGCTTTCCAGTGTCAAGAAAGAAGGACCCATCATCAT
CCACACTGATGAAGCAGATTCAAGACTTGTATCCAACTACCAAAGCTGGAGCCTGA
GGCAGAGAACCGAGGCCGGAGGCAGACTGCCTCTTACAGCCAGGAATCTCAGAGGATTG
AAAAAGGTGAAGGACAGGATGGCATTGACAGTAGTGTAAAGTGGACTTCTCATCCTCCT
GGACAACGTGGCTGCCGAGCAGGCACACAACTCCAAGCTGCCCATGCTGAAGAGATTG
CACGGATGATCGAACAGAGAGCTGTGGACACATCCTGTACATACTGCCAAGGAAGACAGG
GAAAGTCTCAGATGGCAGTAGGCCATTCCACATCCTAGAGAGAACCTGCTGAAAGC
CATGGACTCTGCCACTGCCAACAGATCAGAAAGCTGTATCTATGCCGCTCATGATG
TGACCTTCATACCGCTTTAATGACCTGGGGATTTGACCACAAATGGCCACCGTTGCT
GTTGACCTGACCATGGAACCTTACCAAGCACCTGGAATCTAAGGAGTGGTTGTGCAGCTCTA
TTACACGGGAAGGAGCAGGTGCCAGAGGGTTGCCCTGATGGGCTCTGCCGCTGGACATGT
TCCTGAATGCCATGTCAGTTATACCTTAAGCCCAGAAAAATACCATGCACTCTGCTCTCAA
ACTCAGGTGATGGAAGTTGGAAATGAAGAGTAACTGATTATAAAAGCAGGATGTGTTGATT
TTAAAATAAAGTGCCTTATACAATG

104/246

FIGURE 102

MITGVFSMRLWTPVGVLTSLAYCLHQRRVALAEIQLQEADGQCPVDRSLLKLMVQVVFRHGARSPLKPLPLEEQV
EWPQLLEVPPQTQFDYTVNLAGGPKPSPYDSQYHETTLKGGMFAGQLTKVGMQQMFALGERLRKNYVEDIP
FLSPTFNPQEVFIRSTNIFRNLESTRCLLAGLFQCQKEGPIIIHTDEADSEVLYPNYQSCWSLRQRTRGRRQTA
SLQPGISEDLKKVKDRMGIDSSDKVDFFFILLDNVAEEQAHNLPSCPMLKRFARMIEQRAVDTSLYILPKEDRES
LQMAVGPFHLILESNLKAMDSATAPDKIRKLYIAAHDVTIFIPLIMTLGIFDHKWPPFAVDLTMELYQHLESK
EWFVQLYYHGKEQVPRGCPDGLCPLDMFLNAMSVTLSPEKYHALCSQTQVMEVGNEE

Important features:

Signal sequence:
amino acids 1-23

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 218-222

Casein kinase II phosphorylation site.
amino acids 87-91, 104-108, 320-324

Tyrosine kinase phosphorylation site.
amino acids 280-288

N-myristoylation site.
amino acids 15-21, 117-123, 118-124, 179-185, 240-246, 387-393

Amidation site.
amino acids 216-220

Leucine zipper pattern.
amino acids 10-32

Histidine acid phosphatases phosphohistidine signature.
amino acids 50-65

105/246

FIGURE 103

GGGGCGGGTGGACGCGGACTCGAACGCAAGCTGGCTTCGGGACCCAGGACCCCTCGGGCCCGA
CCCGCCAGGAAAGACTGAGGCCGCGCCTGCCCGCCGGCTCCCTGCGCCGCCGCCCTC
CCGGGACAGAAGATGTGCTCCAGGGCTCCCTGCTGCTGCCGCTGCTCCTGCTACTGCCCT
GGGGCTGGGTGCAAGGGCTGCCATCCGGCTGCCAGTGCAGGCCACAGACAGTCTCT
GCACTGCCGCCAGGGGACCACGGTGCCCGAGACGTGCCACCCGACACGGTGGGCTGTAC
GTCTTGAGAACGGCATACCATGCTGACGCAAGCAGCTTGCCGGCTGCCGGGCTGCA
GCTCTGGACCTGTACAGAACAGATGCCAGCCTGCCCTGCCCTGCTGCTGG
ACCTCAGCCACAACAGCCTGGCCCTGGAGGCCGGCATCTGGACACTGCCAACGTGGAG
GCGCTGCCGCTGGCTGGTCTGGGCTGCAGCAGCTGGACGAGGGCTCTCAGCCGCTTGCG
CAACCTCACGACCTGGATGTGTCGACAACCAGCTGGAGCGAGTGCCACCTGTGATCCGAG
GCCTCCGGGGCCTGACGCGCCTGCCGCTGCCGGCAACACCCGCAATTGCCAGCTGCC
GAGGACCTGCCGGCCTGGCTGCCCTGCCAGGAGCTGGATGTGAGCAACCTAACGCTGCC
CCTGCCTGGCAGCCTCTGCCCTCTTCCCCGCCCTGCCGCTGCTGGCAGCTGCCCAACC
CCTTCAACTGCGTGTGCCCTGAGCTGTTGGCCCTGGTGCAGGAGAGCCACGTCACA
CTGGCCAGCCCTGAGGAGACGCGCTGCCACTTCCCCTCCAGGACTGTAGGGCC
GGAGCTTGAACCGCCGACTTGGCTGCCAGCCACCACACCACAGGCCACAGTGCCCA
CGAGGCCCGGGTGGTGCAGGAGACGCGCTGCCACTTCCCCTCCAGGACTGTAGGGCC
CCCACAGGCCGGCCACTGAGGCCCCAGCCGCCCTCCAGGACTGTAGGGCC
TGTCCCCAGCCCCAGGACTGCCAACCGTCCACCTGCCCTCAATGGGGCACATGCCACCTGG
GGACACGGCACCACTGGCGTGTGGCCCGAAGGGCTCACGGGCTGTACTGTGAGAGC
CAGATGGGGCAGGGGACACGGCCAGCCCTACACCAGTCAGCCGAGGCCACACGGTCC
GACCCCTGGGATCGAGCCGGTGAGCCCCACCTCCCTGCCGTGGGCTGCAAGCCTACCTCC
AGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTATCGAACCTATCGGCC
AAGCGGCTGGTGAACGCTGCACGCTGCCCTCGCTGCTGAGTACACGGTCACCCAGCTGCG
GCCAACGCCACTTACTCCGTCTGTGTCATGCCCTGGGCCCCGGGTGCCGGAGGGCG
AGGAGGCCTGCCGGGAGGCCATACACCCCCAGCCGTCCACTCCAACCACGCCAGTCACC
CAGGCCCGAGGGCAACCTGCCCTCCATTGCCCGCCCTGCCCGGTGCTCTGG
CGCGCTGGCTGCCGTGGGGCAGCCTACTGTGTCGGCGGGGCGGGCATGGCAGCAGCGG
CTCAGGACAAAGGGCAGGTGGGGCAGGGCTGGGCCCCCTGGAACCTGGAGGGAGTGAAGGTC
CCCTTGGAGCCAGGCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGGGGTCTGA
GTGTGAGGTGCCACTCATGGGCTTCCCAGGGCTGCCAGTCACCCCTCCACGCCAAAGC
CCTACATTAAGGCAAGAGAGAGACAGGGCAGCTGGGCGGCTCTCAGCCAGTGAGATGGC
CAGCCCCCTCTGCTGCCACACCACGTAAGTCTCAGTCCCAACCTCGGGATGTGAG
CAGGGCTGTGACGACAGCTGGCCCTGTTCCCTGCCACCTGGCTCTCATCTGTGAG
ATGCTGTGGCCAGCTGACGAGCCCTAACGTCCTCCAGAACCGAGTGCCTATGAGGACAGTGT
CCGCCCTGCCCTCCGCAACGTGCAAGTCCCTGGCACGGCGGGCCCTGCCATGTGCTGTAAC
GCATGCCCTGGCCCTGCTGGCTCTCCACTCCAGGCCACCTGGGGCCAGTGAAGGAAG
CTCCCGAAAGAGCAGAGGGAGAGCGGGTAGGCGGCTGTGACTCTAGTCTTGGCC
AAGCGAAGGAACAAAGAAACTGGAAGGAAGATGCTTAGGAACATGTTGCTTTAA
AATATATATATTTATAAGAGATCCTTCCCATTATCTGGGAAGATGTTTCAAAC
AGAGACAAGGACTTTGGTTTGTAAAGACAAACGATGATATGAAGGCC
ATAAAAAAAAAA

106/246

FIGURE 104

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44804
<subunit 1 of 1, 598 aa, 1 stop
<MW: 63030, pI: 7.24, NX(S/T): 3
MCSRVPLLLPLLLLALGPGVQGCPGQCQCSQPQTVFCTARQGTTVPRDVPPDTVGLYVFEN
GITMLDASSFAGLPGLQLLDLISQNQIASLRLPRLLLLDSHNSLLALEPGILDANVEALRL
AGLGLQQLDEGLFSRLRNHLHLDVSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDLA
GLAALQELDVSNLSQLALPGDLSGLFPRRLLLAAARNPFCNCVCPPLSWFGPWVRESHVTLASP
EETRCHFPPKNAGRLLLELDYADFGCPATTTATVPTTRPVVREPTALSSSLAPTWLSPTAP
ATEAPSPPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGTRHHLAACLCPEGFTGLYCESQMGQ
GTRPSPTPVTPRPPRSLLTGIEPVSPTSLRVGLQRYLQGSSVQLRSRLTYRNLSGPDKRLV
TLRLPASLAEYTVTQLRPNATYSVCVMPGLPGPGRVPEGEEACGEAHTPPAVHSNAPVTQARE
GNLPLLIAPALAAVLLAALAAVGAAYCVRRGRAMAAAAQDKGQVGPAGPLELEGVKVPLEP
GPKATEGGGEALPSGSECEVPLMGFPGLQSPLHAKPYI
```

Signal sequence.

amino acids 1-23

Transmembrane domain.

amino acids 501-522

N-glycosylation sites.

amino acids 198-202, 425-429, 453-457

Tyrosine kinase phosphorylation site.

amino acids 262-270

N-myristoylation sites.

amino acids 23-29, 27-33, 112-118, 273-279, 519-525, 565-571

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

EGF-like domain cysteine pattern signature.

amino acids 355-367

Leucine zipper pattern.

amino acids 122-144, 194-216

107/246

FIGURE 105

CCACACGTCCGAAGGCAGACAAAGGTTCAATTGTAAGAAGCTCCTCCAGCACCTCCTCT
CTTCTCCTTTGCCAAACTCACCCAGTGAGTGTGAGCATTAAAGAACATCCTCTGCCAAG
ACCAAAAGGAAAGAAGAAAAAGGGCCAAAGCCAAATGAAACTGATGGTACTTGTTCAC
CATTGGGCTAACTTTGCTGCTAGGAGTTCAAGCCATGCCCTGCAAATGCCCTCTTGCTACA
GAAAGATACTAAAGATCACAACACTGTCAACACCTCCGGAAGGGAGTAGCTGACCTGACACAG
ATTGATGTCAATGTCAGGATCATTCCTGGATGGGAAGGGATGTGAGATGATCTGTTACTG
CAACTTCAGCGAATTGCTCTGCTGCCAAAGACGTTTCTTGACCAAGATCTCTTCG
TGATTCCCTGCAACAATCATGAAATCTCATGTATTCTGGAGAACACCATTCCCTGATTTC
CCACAAACTGCACTACATCAGTATAACTGCATTCTAGTTCTATATAGTGCATAGAGCAT
AGATTCTATAAATTCTACTTGTCTAAGACAAGTAAATCTGTGTTAACACAAGTAGTAATAAA
AGTTAATTCAATCTAAAAAAAAAAAAAA

108/246

FIGURE 106

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52758
<subunit 1 of 1, 98 aa, 1 stop
<MW: 11081, pI: 6.68, NX(S/T): 1
MKLMVLVFTIGLTLGGVQAMPANRLSCYRKILKDHNCNLPEGVADLTQIDVNQDHFW
DGKGCEMICYCNFSELLCCPKDVFVGPKISFVIPCNNQ

Important features:

Signal peptide:

Amino acids 1-20

N-glycosylation site:

Amino acids 72-76

Tyrosine kinase phosphorylation site:

Amino acids 63-71

109/246

FIGURE 107

AGTGACTGCAGCCTTCCTAGATCCCCTCCACTCGGTTTCTCTCTTGCAGGAGCACCGGCAG
CACCAAGTGTGAGGGGAGCAGGCAGCGGTCTAGCCAGTCCCTGATCCTGCCAGACCACC
CAGCCCCCGGCACAGAGCTGCTCCACAGGCACCATGAGGATCATGCTGCTATTACAGCCAT
CCTGGCCTTCAGCCTAGCTCAGAGCTTGGGGCTGCTGTAAGGAGCCACAGGAGGAGGTGG
TTCCTGGCGGGGGCCGCAGCAAGAGGGATCCAGATCTCTACCAGCTGCTCCAGAGACTCTC
AAAAGCCACTCATCTCTGGAGGGATTGCTCAAAGCCCTGAGGCCAGGCTAGCACAGATCCTAA
GGAATCAACATCTCCCGAGAAACGTGACATGCATGACTTCTTGTGGACTTATGGGCAAGA
GGAGCGTCCAGCCAGAGGGAAAGACAGGACCTTCTTACCTTCAGTGAGGGTTCCCTCGGCC
CTTCATCCCAATCAGCTTGGATCCACAGGAAAGTCTCCCTGGGAACAGAGGAGCAGAGACC
TTTTAAGACTCTCCTACGGATGTGAATCAAGAGAACGTCCCCAGCTTGGCATCCTCAAGTA
TCCCCCGAGAGCAGAATAGGTACTCCACTTCCGGACTCCTGGACTGCATTAGGAAGACCTCT
TTCCCTGTCCCAATCCCCAGGTGCGCACGCTCCTGTTACCCCTTCTCTCCCTGTTCTTGT
ACATTCTTGTGCTTGACTCCTCTCCATCTTCTACCTGACCCCTGGTGAAAAGTGCAT
AGTGAATATCCCCAACCCCAATGGGCATTGACTGTAGAATACCCCTAGAGTTCTGTAGTGT
CTACATTAATAATGTCTCTCTATTCCCTCAACAATAAAGGATTTTGATATGAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

110/246

FIGURE 108

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59849
<subunit 1 of 1, 135 aa, 1 stop
<MW: 14833, pI: 9.78, NX(S/T): 0
MRIMLLFTAILAFSLAQSFAGAVCKEPQEEVVPGGGRSKRDPDLYQLLQRLFKSHSSLEGL
LKALSQASTDPKESTSPEKRDMDHFFVGLMGKRSVQPEGKTGPFLPSVRVPRPLHPNQLG
STGKSSLGTEEQRPL
```

Important features:**Signal peptide:**

Amino acids 1-18

Tyrosine kinase phosphorylation site:

Amino acids 36-45

N-myristoylation sites:

Amino acids 33-39;59-65

Amidation site:

Amino acids 90-94

Leucine zipper pattern:

Amino acids 43-65

Tachykinin family signature:

Amino acids 86-92

111/246

FIGURE 109

GGGGCCACACGCAGCTAGCCGGAGCCCCGACCAGGCGCCTGTGCCTCCTCGTCCCTCGC
CGCGTCCCGCGAAGCCTGGAGCCGGAGCCCCCGCTGCCA**ATG**TGGCGAGCTCAGCA
ACAGGTTCCAAGGAGGGAAAGGCCTCGGCTTGTCAAAGCCGGAGGAGAGGAGGCTGGCC
GAGATCAACCAGGGAGTTCTGTGACAGAAGTACAGTGATGAAGAGAACCTCCAGAAAA
GCTCACAGCCTCAAAGAGAAGTACATGGAGTTGACCTGAACAATGAAGGCGAGATTGACC
TGATGTCTTAAAGAGGATGATGGAGAAGCTTGGTGTCCCCAAGACCCACCTGGAGATGAAG
AAGATGATCTCAGAGGTGACAGGAGGGGTCACTGACACTATATCCTACCGAGACTTGTGAA
CATGATGCTGGGAAACGGTCGGCTGTCTCAAGTTAGTCATGATGTTGAAGGAAAAGCCA
ACGAGAGCAGCCCCAAGCCAGTTGGCCCCCTCCAGAGAGAGACATTGCTAGCCTGCC**GA**
GGACCCCGCCTGGACTCCCCAGCCTCCACCCATACTCCCTCCGATCTGCTGCCCTT
CTTGACACACTGTGATCTCTCTCTCATTGTTGGTCATTGAGGGTTGTTGTGTT
TCATCAATGTCTTGAAAGCACAATTATCTGCCCTAAAGGGGCTCTGGTCGGGAATCC
TGAGCCTTGGGTCCCCCTCCCTCTCTCCCTCCCTCCCCGCTCCCTGTGCAGAAGGGCTG
ATATCAAACCAAAACTAGAGGGGGCAGGGCCAGGGCAGGGAGGCTTCCAGCCTGTGTTCCC
CTCACTTGGAGGAACCAGCACTCTCCATCCTTCAGAAAGTCTCCAAGCCAAGTTCAGGCTC
ACTGACCTGGCTCTGACGAGGACCCAGGCCACTCTGAGAAGACCTGGAGTAGGGACAAGG
CTGCAGGGCCTTTGGGTTCTGGACAGTGCCTGGCCATGGTCCAGTGTCTGGTGTACCC
AGGACACAGCCACTCGGGGCCCCGCTGCCAGCTGATCCCCACTCATTCCACACCTCTTCT
CATCCTCAGTGTGAGGTGGAAAGGAAAGGAGCTTGGCATTGGGAGGCCCTCAAGAAGG
TACCAGAAGGAACCCCTCAGTCTGCTCTGGCCACACCTGTGCAGGCAGCTGAGAGGCAG
CGTGCAGCCCTACTGTCCCTACTGGGGCAGCAGAGGGCTCGGAGGCAGAAGTGAGGCCTG
GGTTTGGGGGGAAAGGTCAAGCTCAGTGTGTTCCACCTTGTAGGGAGGATACTGAGGGGAC
CAGGATGGGAGAATGAGGAGTAAATGCTCACGGCAAAGTCAGCAGCACTGGTAAGCCAAGA
CTGAGAAATACAAGGTTGTTGTCTGACCCCAATCTGCTGAAAAA

112/246

FIGURE 110

MSGELSNRFQGGKAFGLLKARQERRLAEINREFLCDQKYSDEENLPEKLTAFKEKYMEDLN
NEGEIDLMSLKRMMEKLGVPKTHLEMKKMISEVTGGVSDTISYRDFVNMMMLGKRSAVLKVM
MFEKGKANESSPKPVGPPPEDIASLP

FIGURE 111

TAAAACAGCTACAATATTCCAGGGCCAGTCACTTGCCATTCTCATACAGCGTCAGAGAGA
AAGAACTGACTGAAACGTTGAGATGAAGAAAGTCTCCTCCTGATCACAGCCATCTGGCA
GTGGCTGTTGGTTCCCAAGTCTCAAGACCAGGAACGAGAAAAAGAAGTATCAGTGACAG
CGATGAATTAGCTTCAGGGTTTTGTGTTCCCTACCCATATCCATTTCGCCCCACTTCCAC
CAATTCCATTCCAAGATTCATGGTTAGACGTAATTTCCTATTCCAATACCTGAATCT
GCCCTACAACCTCCCTCTAGCGAAAAGTAAACAAGAAGGATAAGTCACGATAAACCTGG
TCACCTGAAATTGAAATTGAGCCACTTCCTGAAGAATCAAATTCCGTAAATAAAAGAAA
AACAAATGTAATTGAAATAGCACACAGCATTCTCTAGTCATATCTTAGTGATCTTCTTA
ATAAACATGAAAGCAAAGATTGGTTCTTAATTCCACA

114/246

FIGURE 112

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71290
><subunit 1 of 1, 85 aa, 1 stop
><MW: 9700, pI: 9.55, NX(S/T): 0
MKKVLLLITAILAVAVGFPVSDQEREKRSISDSDELASGFFVFPYPPFRPLPPIPFP
FPWFRRNFPPIPESAPTTPLPSEK
```

Important features of the protein:

Signal peptide:

Amino acids 1-17

Homologous region to B3-hordein:

Amino acids 47-85

115/246

FIGURE 113

CTCCTCTAACATACTTGCAGCTAAACTAAATATTGCTGCTGGGGACCTCCTCTAGCCT
TAAATTCAGCTCATCACCTCACCTGCCTGGTCATGGCCTGCTATTCTCCTGATCCT
GCCATTGCAACCAGACCTGGATTCTAGCGTCTCCATCTGGAGTGCGGCTGGTGGGGGCCT
CCACCGCTGTGAAGGGCGGGTGGAGGTGGAACAGAAAAGGCCAGTGGGGCACC GTGTGATG
ACGGCTGGGACATTAAGGACGTGGCTGTGTTGCCGGGAGCTGGCTGTGGAGCTGCCAGC
GGAACCCCTAGTGGTATTTGTATGAGCCACCAGCAGAAAAAGAGCAAAGGTCCCTCATCCA
ATCAGTCAGTTGCACAGGAACAGAAAGATACATTGGCTCAGTGTGAGCAAGAAGAAGTTATG
ATTGTTCACATGATGAAGATGCTGGGCATCGTGTGAGAACCCAGAGAGCTCTTCTCCCCA
GTCCCAGAGGGTGTCAAGGCTGGCTGACGGCCCTGGGCATTGCAAGGGACGCGTGGAAAGTGAA
GCACCAGAACCAAGTGGTATACCGTGTGCCAGACAGGCTGGAGCCTCCGGCCGCAAAGGTGG
TGTGCCGGCAGCTGGGATGTGGGAGGGCTGTACTGACTCAAAACGCTGCAACAAGCATGCC
TATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCTAGCTCAGGACGAGAACCCCTTCA
GGATTGCCCTCTGGCCTTGGGGAAAGAACACCTGCAACCATGATGAAGACACGTGGTCG
AATGTGAAGATCCCTTGACTTGAGACTAGTAGGAGGAGACAACCTCTGCTCTGGCGACTG
GAGGTGCTGCACAAGGGCGTATGGGCTCTGTCTGTGATGACAACACTGGGAGAAAAGGAGGA
CCAGGTGGTATGCAAGCAACTGGCTGTGGGAAGTCCCTCTCCCTCAGAGACCGGA
AATGCTATGCCCTGGGTTGCCGCATCTGGCTGGATAATGTTGCTGCTCAGGGAGGGAG
CAGTCCCTGGAGCAGTGCACAGATTGGGGTTTCAGACTGCACCCACCAGGAAGA
TGTGGCTGTCATCTGCTCAGTGTAGGTGGCATTGATCTAATCTGTTGAGTGCCTGAATAGAA
GAAAAACACAGAAGAAGGGAGCATTACTGTCTACATGACTGCATGGGATGAACACTGATCT
TCTTCTGCCCTGGACTGGGACTTATACTTGTTGCCCTGATTCTCAGGCCTCAGAGTTGG
ATCAGAACTTACAACATCAGGTCTAGTTCTCAGGCCATCAGACATAGTTGGAACATACATCA
CCACCTTCCTATGTCACATTGCACACAGCAGATTCCAGCCTCATAATTGTGTAT
CAACTACTAAATACATTCTCACACACACACACACACACACACACACACACATA
CACCATTGCTGTTCTGAGAAACTCTGACAAAATACAGATTGGTACTGAAAGAGA
TTCTAGAGGAACGGAATTAAAGGATAAATTCTGAATTGGTTATGGGTTCTGAAATTG
GCTCTATAATCTAATTAGATATAAAATTCTGGTAACCTTATTACAATAAAAGATAGCAC
TATGTGTTCAAA

116/246

FIGURE 114

MALLFSLILAICTRPGLASPSGVRLVGLHRCEGRVEVEQKGQWGTVCDDGWDIKDVAVLC
RELGCGAASGTPSGILYEPPAEEKQKVLIQSVSCTGTEDTLAQCEQEEVYDCSHDEDAGASC
ENPESSFSPVPEGVRLADGPGHCKGRVEVKHQNQWYTVQCQGWSLRAAKVVCRCQLGCGRAVL
TQKRCNKHAYGRKPIWLSQMSCSGREATLQDCPSGPWGKNTCNHDEDTWVECEDPFDRLLVG
GDNLCSGRLEVHKGVWGSVCDDNWGEKEDQVVCKQLGCGKSLSPSFRDRKCYGPGVGRIWL
DNVRCSGEEQSLEQCQHRFWGFHDCTHQEDVAVICSV

Signal sequence:
amino acids 1-15

Casein kinase II phosphorylation site.
amino acids 47-51, 97-101, 115-119, 209-213, 214-218, 234-238,
267-271, 294-298, 316-320, 336-340

N-myristoylation site.
amino acids 29-35, 43-49, 66-72, 68-74, 72-78, 98-104, 137-143,
180-186, 263-269, 286-292

Amidation site.
amino acids 196-200

Speract receptor repeated domain signature.
amino acids 29-67, 249-287

117/246

FIGURE 115

CATTCCAACAAGAGCACTGGCCAAGTCAGCTTCTTGAGAGAGTCTAGAAGACATGAT
GCTACACTCAGCTTGGGTCTGCCTCTACTCGTCACAGTTCTCCAACCTGCCATTG
CAATAAAAAGGAAAAGAGGCCTCCTCAGACACTCTCAAGAGGATGGGGAGATGACATCACT
TGGGTACAAACTTATGAAGAAGGTCTTTATGCTCAAAAAAGTAAGAAGCCATTAATGGT
TATTCACTCACCTGGAGGATTGCAATACTCTCAAGCACTAAAGAAAGTATTGCCAAAATG
AAGAAATACAAGAAATGGCTCAGAATAAGTCATCATGCTAAACCTTATGCATGAAACCAC
GATAAGAATTATCACCTGATGGCAATATGTGCTAGAATCATGTTGTAGACCCCTTT
AACAGTTAGAGCTGACATAGCTGGAAGATACTCTAACAGATTGTACACATATGAGCCTCGGG
ATTACCCCTATTGATAGAAAACATGAAGAAAGCATTAAGACTTATTCAAGTCAGAGCTATAA
GAGATGATGGAAAAAGCCTTCACTTCAAAGAAGTCAAATTGATGAAGAAAACCTCTGGCA
CATTGACAAATACTAAATGTGCAAGTATATAGATTTGTAATATTACTATTTAGTTTTTA
ATGTGTTGCAATAGTCTTATTAAAATGTTTTAAATCTGA

118/246

FIGURE 116

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64896
<subunit 1 of 1, 166 aa, 1 stop
<MW: 19171, pI: 8.26, NX(S/T): 1
MMLHSALGLCLLLVTVSSNLIAIAIKKEKRPPQTLSRGWGDDITWVQTYEEGLFYAQKSKK
PLMVIHHLEDCQYSQALKKVFAQNEEIQEMAQNKFIMLNLMHETTDKNLSPDGQYVPRIM
FVDPSLTVRADIAGRYSNRLYTYEPRDPLLLIENMKKALRLIQSEL
```

Important features:**Signal peptide:**

Amino acids 1-23

N-myristoylation site:

Amino acids 51-57

119/246

FIGURE 117

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCAGGCTCCAGGTGGGTCGGTCCGCATCCA
GCCTAGCGTGTCCACGATGCGGCTGGGCTCCGGACTTCGCTACCTGTTGCGTAGCGATCG
AGGTGCTAGGGATCGCGGTCTCCTCGGGATTCTTCCCGGCTCCCGTCTGTTCTCTGCC
AGAGCGGAACACGGAGCGGAGCCCCCAGCGCCCGAACCCCTCGGCTGGAGCCAGTTCTAAGT
GACCACGCTGCCACCACCTCTCAGTAAAGTTGTTATTGTTCTGATAGATGCCCTGAGAG
ATGATTTGTGTTGGTCAAAGGGTGTGAAATTATGCCCTACACAACCTACCTTGTGGAA
AAAGGAGCATCTCACAGTTGTGGCTGAAAGCAAAGCCACCTACAGTTACTATGCCCTCGAAT
CAAGGCATTGATGACGGGGAGCCTCTGGCTTGTGACGTACAGGAACCTCAATTCTC
CTGCACTGCTGGAAGACAGTGTGATAAGACAAGCAAAGCAGCTGGAAAAAGAATAGTCTT
TATGGAGATGAAACCTGGGTTAAATTATCCAAAGCATTGTGGAATATGATGGAACAAAC
CTCATTTCGTGTCAGATTACACAGAGGTGGATAATAATGTCACGAGGCATTGGATAAAG
TATTAAGGAGAGATTGGGACATATTAATCCTCCACTACCTGGGGCTGGACCACATTGGC
CACATTCAAGGGCCCAACAGCCCCCTGATTGGGAGAAGCTGAGCGAGATGGACAGCGTGCT
GATGAAGATCCACACCTCACTGCAAGGAGAGAGAGACGCCATTACCAATTGCTGG
TTCTTGTGGTGACCATGGCATGTCAGAAACAGGAAGTCACGGGCTCCTCCACCGAGGAG
GTGAATACACCTCTGATTTAATCAGTTCTGCGTTGAAAGGAAACCCGGTGTATCCGACA
TCCAAAGCACGTCATAGACGGATGTGGCTGCGACACTGGCGATAGCAGTTGGCTTACCGA
TTCCAAAAGACAGTGTAGGGAGCCTTATTCCAGTTGTGGAAGGAAGACCAATGAGAGAG
CAGTTGAGATTTACATTGAATACAGTCAGCTTAGTAAACTGTTGCAAGAGAATGTGCC
GTCATATGAAAAGATCCTGGTTTGAGCAGTTAAAATGTCAGAAAGATTGCAATGGGAAC
GGATCAGACTGACTGGAGGAAAGCATTCAAGACTGCTTATTCAACCTGGCTCCAAGGTT
CTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTTGTCCCTGAGTGCACAAGTGGCCA
GTTCTACCCCTGCTCCTGCTCAGCGTCCACAGGCAGTCACAGAAAGGCTGAGCTGGAAAGTC
CCACTGTCATCTCTGGTTCTGCTCTTTATTGGTGTGATCCTGGTTCTTCGGCCGT
TCACGTCATTGTGTCACCTCAGCTGAAAGTTCTGCTACTTCTGTGGCTCTCGTGGCTGG
CGGCAGGCTGCGCTTCGTTACAGACTCTGGTGAACACCTGGTGTGCAAGTGTGG
AGTGCCTGGACAGGGGGCCTCAGGGAGGACGTGGAGCAGCCTTATCCAGGCCTCTGGGT
GTCCCGACACAGGTGTCACATCTGTGCTGTCAGGTAGATGCCCTAGTTCTGGAAAGCTA
GGTTCTGCACTGTTACCAAGGTGATTGTAAGAGCTGGCGGTACAGAGGAACAAGCCCC
CCAGCTGAGGGGGTGTGTGAATCGGACAGCCTCCAGCAGAGGTGTGGAGCTGCAGTGAG
GGAAGAAGAGACAATCGGCCTGGACACTCAGGAGGGTCAAAGGAGACTTGGTCGCACCACT
CATCCTGCCACCCCCAGAATGCATCCTGCCATCAGGTCCAGATTCTTCCAAGGCGGAC
GTTTCTGTTGAAATTCTTAGTCCTGGCCTGGACACCTTCATTGTTAGCTGGGAGTGG
TGGTGAGGCAGTGAAGAAGAGGCAGGACTGTTGGCCCCCACCCCAACCTGCACAGCCCTCATCC
CCTCTGGCTTGAGCGTCAGAGGCCCTGTGCTGAGTGTCTGACCGAGACACTCACAGCTT
GTCATCAGGGCACAGGCTCCTCGGAGGCCAGGATGATCTGTCGCCACGCTTGACACTCGGGCC
CATCTGGGCTCATGCTCTCTCTGCTATTGAATTAGTACCTAGCTGCACACAGTATGTAG
TTACCAAAAGAATAACGGCAATAATTGAGAAAAAAA

120/246

FIGURE 118

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA84920
><subunit 1 of 1, 310 aa, 1 stop
><MW: 33875, pI: 7.08, NX(S/T): 2
MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTTL
PPPLFSKVIVLVIDALRDDFVFGSKGVKFMPYTTYLVEKGASHSFVAEAKPPTVTMPRIK
ALMTGSILPGFVDVIRNLNSPALLEDSVIRQAKAAGKRIVFYGDETWVKLFPKHFVEYDGT
TSFFVSDYTEVDNNVTRHLDKVLKRGDWLILHYLGLDHIGHISGPNSPLIGQKLSEMD
SVLMKIHTSLQSKERETPLPNLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKP
GDIRHPKHVQ
```

Important features of the protein:**Signal peptide:**

Amino acids 1-34

Transmembrane domain:

Amino acids 58-76

N-glycosylation sites:

Amino acids 56-60;194-198

N-myristoylation sites:

Amino acids 6-12;52-58;100-106;125-131;233-239;270-276;
275-281;278-284

Amidation site:

Amino acids 154-158

Cell attachment sequence:

Amino acids 205-208

121/246

FIGURE 119

GCCCACGCGTCCGATGGCGTTACGTTCGCGGCCCTCTGCTACATGCTGGCGCTGCTGCTCA
CTGCCCGCCTCATCTTCTGCCATTGGCACATTAGCATTTGATGAGCTGAAGACTGAT
TACAAGAATCCTATAGACCAGTGTAAATACCCCTGAATCCCCTGTACTCCCAGAGTACCTCAT
CCACGCTTCTTCTGTGTATGTTCTTGTGCAGCAGAGTGGCTTACACTGGGTCTCAATA
TGCCCCCTCTGGCATATCATATTGGAGGTATATGAGTAGACCAGTGTAGTGGCCAGGA
CTCTATGACCCCTACAACCATCATGAATGCAGATATTCTAGCATATTGTCAAAGGAAGGATGG
TGCAAATTAGCTTTTATCTCTAGCATTTTACTACCTATATGGCATGATCTATGTTT
GGTAGCTCTTAGAACACACAGAAGAATTGGTCCAGTTAAGTGCATGCAAAAAGCCACC
AAATGAAGGGATTCTATCCAGCAAGATCCTGTCCAAGAGTAGCCTGTGGAATCTGATCAGTT
ACTTTAAAAAATGACTCCTTATTTTAAATGTTCCACATTTGCTTGTGAAAGACTGT
TTTCATATGTTATACTCAGATAAAGATTAAATGGTATTACGTATAAATTAAATATAAAATG
ATTACCTCTGGTGTGACAGGTTGAACCTGCACCTCTTAAGGAACAGCCATAATCCTCTGA
ATGATGCATTAATTACTGACTGCTCTAGTACATTGGAAAGCTTTGTTATAGGAACCTGTAG
GGCTCATTGGTTTCATTGAAACAGTATCTAATTAAATTAGCTGTAGATATCAGGTGCT
TCTGATGAAGTGAATGTATATCTGACTAGTGGAAACTTCATGGGTTCCATCTGTCA
TGTGATGATTATATGGATACATTACAAAAAATGGGAATTTCCTCGCTTG
AATATTATCCCTGTATATTGCATGAATGAGAGATTCCATATTCCATCAGAGTAATAAAT
ATACTTGCTTAATTCTAACGATAAGTAAACATGATATAAAATATGCTGAATTACTTG
TGAAGAATGCATTTAAAGCTATTAAATGTGTTTATTTGTAAGACATTACTTATTAAGA
AATTGGTTATTATGCTTACTGTTCTAATCTGGTGGTAAAGGTATTCTAAGAATTGCAGGT
ACTACAGATTTCAAAAGTGAATGAGAGAAAATTGTATAACCCTGCTGTTCTTAGTG
CAATACAATAAAACTCTGAAATTAAAGACTC

122/246

FIGURE 120

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23330
<subunit 1 of 1, 144 aa, 1 stop
<MW: 16699, pI: 5.60, NX(S/T): 0
MAFTFAAFCYMLALLTAALIFFAIWHIAFDELKTDYKNPIDQCNILNPLVLPEYLIHA
FFCVMFLCAAELTLGLNMPLLAYHIWRYMSRPVMSGPGLYDPTTIMNADILAYCQKEGW
CKLAFYLLAFFYYLYGMIYVLVSS

Important features:

Signal peptide:

Amino acids 1-20

Type II transmembrane domain:

Amino acids 11-31

Other transmembrane domain:

Amino acids 57-77;123-143

Glycosaminoglycan attachment site:

Amino acids 96-100

123/246

FIGURE 121

CGGACGCGTGGCGGACGCGTGGCGGCCACGGCGCCCGGGCTGGGCGGTGCTTCTT
CCTTCTCCGTGGCCTACGAGGGTCCCCAGCCTGGTAAAGATGGCCCCATGGCCCCCGAAGG
GCCTAGTCCCAGCTGTGCTCTGGGCCTCAGCCTTCTCCTCAACCTCCCAGGACCTATCTGG
CTCCAGCCCTCTCCACCTCCCCAGTCTTCTCCCCCGCTCAGCCCCATCCGTGTACACCTG
CCGGGGACTGGTACAGCTTAACAAGGGCTGGAGAGAACCATCCGGACAACACTTGGAG
GTGGAAACACTGCCTGGGAGGAAGAGAACATTGTCAAATACAAAGACAGTGAGACCCGCCTG
GTAGAGGTGCTGGAGGGTGTGCAAGTCAGACTTCGAGTGCCACCGCCTGCTGGAGCT
GAGTGAGGAGCTGGTGGAGAGCTGGTTACAAGCAGCAGGAGGGCCCGAACCTCTTCC
AGTGGCTGTGCTCAGATTCCCTGAAGCTCTGCTGCCCGCAGGCACCTCGGGCCCTCCTGC
CTTCCCTGTCTGGGGAAACAGAGAGGCCCTGCGGTGGCTACGGGCAGTGTGAAGGAGAAGG
GACACGAGGGGGCAGCGGGCACTGTGACTGCCAAGCCGGCTACGGGGGTGAGGCCTGTGGCC
AGTGTGGCCTTGGCTACTTGTGAGGCAGAACGCAACGCCAGCCATCTGGTATGTTGGCTTGT
TTTGGCCCTGTGCCGATGCTCAGGACCTGAGGAATCAAACGTGTTGCAATGCAAGAAGGG
CTGGGCCCTGCATCACCTCAAGTGTGAGACATTGATGAGTGTGGCACAGAGGGAGCCAAGT
GTGGAGCTGACCAATTCTCGTGAACACTGAGGGCTCTATGAGTGCCGAGACTGTGCCAAG
GCCTGCCTAGGCTGCATGGGGCAGGGCAGGTCGCTGTAAGAAGTGTAGCCCTGGCTATCA
GCAGGTGGCTCCAAGTGTCTCGATGTGAGTGTGAGACAGAGGTGTGTCGGAGAGAAGA
ACAAGCAGTGTGAAAACACCGAGGGCGTTATCGCTGCATCTGTGCCGAGGGCTACAAGCAG
ATGGAAGGCATCTGTGTGAAGGGAGCAGATCCCAGAGTCAGCAGGCTTCTCAGAGATGAC
AGAAGACGAGTTGGTGGTGTGCAAGCAGATGTTCTTGGCATCATCATCTGTGCACTGGCCA
CGCTGGCTGCTAAGGGCAGCTGGTGTTCACCGCCATCTCATTGGGGCTGTGGCGGCCATG
ACTGGCTACTGGTTGTCAAGAGCGAGCTGACCGTGTGCTGGAGGGCTTCATCAAGGGCAGATA
ATCGCGGCCACCACCTGTAGGACCTCTCCACCCACGCTGCCCGAGCTGGGCTGCC
TCCGTGGACACTCAGGACAGCTGGTTATTTGAGAGTGGGTAAGCACCCCTACCTG
CCTTACAGAGCAGCCAGGTACCCAGGCCGGCAGACAAGGCCCTGGGTAAAAAGTAGC
CCTGAAGGTGGATACCATGAGCTTCACTGGCGGGACTGGCAGGCTCACAAATGTGTGA
ATTCAAAAGTTTCTTAATGGTGGCTGCTAGAGCTTGGCCCTGCTTAGGATTAGGTG
GTCCTCACAGGGTGGGCCATCACAGCTCCCTCTGCCAGCTGCATGTCAGTGGCT
TCTGTGTTACCAACATCCCCACACCCATTGCCACTTATTATTCATCTCAGGAAATAAGA
AAGGTCTTGGAAAGTAAAAAAAAAAAAAAAAAAAAAA

124/246

FIGURE 122

MAPWPPKGVLVPAVLWGLSLFLNLPGPIWLQPSPPQSSPPPQPHPCHTCRGLVDSFNKGLER
TIRDNFGGGNTAWEENLSKYKDSETRLVEVLEGVCSKSDFECHRLLELSEELVESWWFHKQ
QEAPDLFQWLCSDSLKLCCPAGTFGPSCLPAGTERPCGGYGQCEGEGTRGGSGHCDCQAG
YGEACGQCGLGYFEAERNASHLVCASFAGPCARCSGPEESNCLQCKKGWALHHLCVDIDE
CGTEGANCGADQFCVNTESYECRDCAKACLGCMAGPGRCKCSPGYQQVGSKCLDVDECE
TEVCPGENKQCENTEGGYRCICAEGYKQMEGICVKEQIPESAGFFSEMTEDELVVLQQMFFG
IIICALATLAAKGDLVFTAIFIGAVAAMTGYWLRSERSDRVLEGFIKGR

Important features:**Signal peptide:**

Amino acids 1-29

Transmembrane domain:

Amino acids 342-392

N-glycosylation sites:

Amino acids 79-83;205-209

cAMP- and cGMP-dependent protein kinase phosphorylation site:
Amino acids 290-294**Aspartic acid and asparagine hydroxylation site:**

Amino acids 321-333

EGF-like domain cysteine pattern signature:

Amino acids 181-193

125/246

FIGURE 123

GCAAGCCAAGGCGCTGTTGAGAAGGTGAAGAAGTCCGGACCCATGTGGAGGAGGGGACATTGTGTACGCC
TCTACATGCGCAGACCATCATCAAGGTGATCAAGTTCATCCTCATCATCTGCTACACCGTCTACTACGTGCAC
AACATCAAGTTCGACGTGACTGCACCGTGGACATTGAGAGCCTGACGGGCTACCGCACCTACCGCTGTGCCA
CCCCCTGGCCACACTCTCAAGATCCTGGCCTCTACATCAGCCTAGTCATCTCATCTACGGCCTCATCTGCA
TGTACACACTGTGGGGATGTCAGGGCCTCCCTCAAGAAGTACTCGTTGAGTCAGTGCACCTCATGCTCA
TACAGCGACATCCCCGACGTCAAGAACGACTTCGGCTTCATGCTGCACCTCATGACCAATACGACCCGCTTA
CTCCAAGCGCTTCGCCGCTTCCCTGCGGAGGTGAGTGGAGAACAGCTGCCAGCTGAACCTCAACAAACGAGT
GGAGCCTGGACAAGCTCCGGCAGCGCTACCAAGAACGCGCAGGACAAGCTGGAGCTGCACCTGTTCATGCTC
AGTGGCATCCCTGACACTGTGTTGACCTGGTGGAGCTGGAGGTCTCAAGCTGGAGCTGATCCCCGACGTGAC
CATCCCGCCAGCATGCCAGCTCACGGGCTCAAGGAGCTGTTGACCTGGTGGAGCTGGAGGTCTACACAGCGGCCAAGATTGAAG
CGCCTGCCGCTGGCCTTCTGCGCAGAACCTGCGGGCCTGACATCAAGTTCACCGACATCAAGGAGATCCCG
CTGAGATCTATAGCCTGAAAGACACTGGAGGAGCTGACCTGACGGCAACCTGAGCGGGAGAACAAACCGCTA
CATCGTCATGCCACGGGCTGCCGGAGCTCAAACGCCCTCAAGGTGCTGCCGCTCAAGAGCAACCTAAGCAAGCTGC
CACAGGTGGTACAGATGTGGCGTGCACCTGCGAGAACAGCTGTCATCAACAATGAGGGCACCAAGCTCATCGTC
CTCAACAGCCTCAAGAAGATGGCAACCTGACTGAGCTGGAGCTGATCCGCTGCGACCTGGAGCGCATCCCCCA
CTCCATCTTCAGCCTCCACAACCTGCGAGGAGATGACCTCAAGGACAAACACCTCAAGGACATCGAGGAGATCA
TCAGCTTCCAGCACCTGACCCGCTCACCTGCCCTAAAGCTGTTGACCAACCACATGCCCTACATCCCCATCCAG
ATCGGCAACCTCACCAACCTGGAGCCCTTACCTGACCCACAACAGCTGACCTTCCCTGCCGACATCGGCCCT
CTACTGCGCAAGCTGCGTACCTGGACCTCAGGCCACAACACCTGACCTTCCCTGCCGACATCGGCCCT
TGCAGAACCTCAGAACCTAGCCATCAGGCCAACGGATGAGACGCTCCCTCCGGAGCTTCCAGTGC
AACGCTGCCGAGGAGCTGCGGGCAACCCGCTGGAGTGCCTGCTGTGGAGCTGGCGAGTGCCTCAAGC
GACCGAGATCGAGCTGCCGGCAACCCGCTGGAGTGCCTGCTGTGGAGCTGGCGAGTGCCTCAAGC
GCAGCGGCTTGGTGGAGGAGGACCTGTTAACACACTGCCACCCGAGGTGAAGGAGCGGCTGTGGAGGGCT
GACAAGGAGCAGGCCTGAGCGAGGCCGGCCAGCACAGCACAGCAGCAGGACCGCTGCCAGTCCAGGCCGG
AGGGGCAGGCCCTAGCTCTCCAGAACCTCCGGACAGCCAGGCCAGCCTGCCGCTGGCGAGGAGCCTGGGCC
GCTGTGAGTCAGGCCAGAGCAGAGGACAGTATCTGTGGGCTGCCCTTTCTCCCTGAGACTCACGTC
CCCCAGGGCAAGTGTGGAGGAGCAAGTCTCAAGAGCGCAGTATTGGATAATCAGGGCTCCTCCCTG
GAGGCCAGCTGCCCCAGGGCTGAGCTGCCACAGGAGCTGGACCTCACTTACTTACTTACTTACTT
TTTCTCCATCTCCACCTCCTCATCCAGATAATTACATTCCAAAGAAAGTTAGCGATGCCGCGGGCATTTAACACCCACCTGG
TTCAAGGGAAAGTGGCTGCCCTTCCCTGTCTTATTAGCGATGCCGCGGGCATTTAACACCCACCTGG
ACTTCAGCAGAGTGGTCCGGCGAACCAGCCATGGGACGGTCAAGGCCAGTGCCTCTCAGTTGTGGAGTT
CGGTCCACGGAGAGCAGGCCCTCAGCTGGAAAGGCCAGGCCAGTGCCTCTCAGTTGTGGAGTT
TTAGTTTTGTTTTTTTAATCAAAAAAAACATTTTAAAAAAAGACACTAACGGCCAGTGAAGTTGGAGTCTCAGGGCAGG
GTGGCAGTTCCCTTGAGCAAAGCAGCCAGACGTTGAACTGTGTTCCCTGGCGCAGGGTGCAGGGTG
TCTCCGGATCTGGTGTGACCTGGCAGGAGTTCTATTGTTCTGGGAGGGAGGTTTTGTGTTTT
TTGGGTTTTTTGGTGTCTGTTCTTCTCTCATGTGCTTGGCAGGCACTCATTTCTGTGGCTGTGGC
CAGAGGGAAATGTTCTGGAGCTGCCAAGGAGGGAGGAGACTCGGGTGGCTAATCCCGGATGAAACGGTGTCCA
TTCCGACCTCCCTCCTCGTGCCTGCCCTGCCCTCCACGCAAGTGTAAAGGAGCCAAGAGGGAGCCACTTCGC
CCAGACTTTGTTCCCCACCTCTGCCGATGGGTGTGCCAGTGCCACCGCTGGCCTCCGCTGCTCCATCAG
CCCTGTCGCCACCTGGTCTTCAAGAGAGCAGACACTTAGAGGCTGGCAGGAAATGGGAGGTGCGCCCTGG
AGGGCAGGCCGGTGGTCCAAGGCCGGTCCCGTCCAGTGGAGTGCACACAGCCAGTGCAGGCCACCTGG
GCTGGAGCCAACCTGCTTAGATCACTGGGCCCCACCTTAGAAGGGTCCCCGCTTAGATCAATCAGTGG
ACACTAAGGCACGTTTAGAGTCTCTGTCTTAATGATTATGTCATGCCCTGTCCTGCA
GCGTCGTGTCATTGGATATAATCCTCAGAAATAATGCACACTAGCCTCTGACAACCATGAAGAAAATCCGTT
ACATGTGGCTGAACCTGTAGACTCGGTACAGTCAAATAACAGAAAAAA

126/246

FIGURE 124

MRQTIIKVIKFILIIICYTVYYVHNIKFVDVDCVDIESLTGYRTYRCAHPLATLFKILASFYI
SLVIFYGLICMYTLWWMLRRSLKKYSFESIREESSYSDIPDVKNDFAFMLHLIDQYDPLYSK
RFAVFLSEVSENLRLQLNLNNEWTLDKLRLQRLTKNAQDKLELHLFMLSGIPDTVFDLVELEV
LKLELIPDVTIPPSIAQLTGLKELWLYHTAAKIEAPALAFIRENLRALHIKFTDIKEIPLWI
YSLKTLEELHLTGNLSAENNRYIVIDGLRELKRLKVLRLKSNLSKLPQVVTDVGVHLQKLSI
NNEGTKLIVLNSLKKMANLTELLELIRCDLERIPHISIFSLHNLQEIDLKDNNLKTIEEIISFQ
HLHRLTCLKLWYNHIAYIPIQIGNLTNERLYLNRNKIEKIPPTQLFYCRKLRYLDLSHNNLT
FLPADIGLILQNLQNLAITANRIETLPPELFQCRKLRALHLGNNVLQSLPSRVGELTNLTQIE
LRGNRLECLPVELGECPLLKRSGLVVEEDLFNTLPPEVKERLWRADKEQA

Transmembrane domain:

amino acids 51-75 (type II)

N-glycosylation site.

amino acids 262-266, 290-294, 328-332, 396-400, 432-436, 491-495

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 85-89

Casein kinase II phosphorylation site.amino acids 91-95, 97-101, 177-181, 253-257, 330-334, 364-368,
398-402, 493-497**N-myristoylation site.**

amino acids 173-179, 261-267, 395-401, 441-447

127/246

FIGURE 125

GTTGTGTCTTCAGCAAAACAGTGGATTAAATCTCCTGCACAAGCTGAGAGCAACACAA
TCTATCAGGAAAGAAAGAAAGAAAAACCGAACCTGACAAAAAAGAAGAAAAAGAAGA
AAAAAAATCATGAAAACCATCCAGCCAAAATGCACAATTCTATCTCTGGGCAATCTCAC
GGGGCTGGCTGCTCTGTCTTCCAAGGGAGTCCCCGTGCGCAGCAGGAGATGCCACCTTC
CCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCTCAGGTGCACTATT
GACAACCGGGTCACCCGGTGGCCTGGCTAAACCGCAGCACCATCCTATGCTGGGAATGA
CAAGTGGTGCCTGGATCCTCGCGTGGCTTCTGAGCAACACCCAAACGAGTACAGCATCG
AGATCCAGAACGTGGATGTGTATGACGAGGGCCCTAACACCTGCTCGTGCAGACAGACAAAC
CACCCAAAGACCTCTAGGTCCACCTCATTGTGCAAGTATCTCCAAAATTGTAGAGATTTC
TTCAGATATCTCCATTAATGAAGGGAACAAATTAGCCTCACCTGCATAGCAACTGGTAGAC
CAGAGCCTACGGTTACTGGAGACACATCTCTCCAAAGCGGTTGGCTTGTGAGTGAAGAC
GAATACTTGGAAATTCAAGGGCATCACCCGGAGCAGTCAGGGACTACGAGTGCAGTGCCTC
CAATGACGTGGCCGCCGTGGTACGGAGAGTAAAGGTACCGTGAACATATCCACCATACA
TTTCAGAAGCCAAGGGTACAGGTGTCCCCGTGGACAAAAGGGGACACTGCAGTGTGAAGCC
TCAGCAGTCCCCTCAGCAGAATTCCAGTGGTACAAGGATGACAAAGACTGATTGAAGGAAA
GAAAGGGGTGAAAGTGGAAAACAGACCTTCTCTCAAAACTCATCTCTTCAATGTCTTG
AACATGACTATGGGAACTACACTTGCGTGGCCTCCAACAAGCTGGCACACCAATGCCAGC
ATCATGCTATTGGTCCAGGCGCCGTAGCGAGGTGAGCAACGGCACGTCGAGGAGGGCAGG
CTGCGTCTGGCTGCTGCCTCTTGTGACCTGCTTCTCAAATTTGATTGTGAGTGCC
ACTTCCCCACCCGGAAAGGCTGCCACCACCAACACAGCAATGGCAACAC
CGACAGCAACCAATCAGATATATAAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGA
AATTGAGGGAGGGAAACAAAGAATACTTGGGGGGAAAAGAGTTAAAAAAAGAAATTGAA
AATTGCCTTGCAGATATTAGGTACAATGGAGTTTCTTCCCAAACGGGAAGAACACAGC
ACACCCGGCTTGGACCCACTGCAAGCTGCATCGTCAACCTCTTGGTGCCAGTGTGGCAA
GGGCTCAGCCTCTGCCCCACAGAGTCCCCCACGTGGAACATTCTGGAGCTGGCCATCCCA
AATTCAATCAGTCATAGAGACGAACAGAATGAGACCTCCGGCCAAAGCGTGGCGTGC
GCACTTGGTAGACTGTGCCACCACGGCGTGTGAAACGTGAAATAAAAGAGCAAAA
AAAAAA

128/246

FIGURE 126

MKTIQPKMHNSISWAIFTGLAALCLFQGVPVRSGDATFPKAMDNVTVRQGESATLRCTIDNR
VTRVAWLNRSTILYAGNDKWCLDPRVLLSNTQTQYSIEIQNVDVYDEGPYTCVQTDNHPK
TSRVHLIVQVSPKIVEISSLSDISINEGNNISLTCIATGRPEPTVTWRHISPKAvgFVSEDEYL
EIQQITREQSGDYECASNDVAAPVVRVKVTVNYPPYISEAKGTGPVGQKGTQCEASAV
PSAEFQWYKDDKRLIEGKKGVKVENRPFLSKLIFNVSEHDYGNYTCVASNKLGHTNASIML
FGPGAVSEVSNGTTSRAGCVWLLPLLVLHLLLKF

Important features:

Signal peptide:

amino acids 1-28

129/246

FIGURE 127

GGCGCCGGTGCACCGGGCGGGCTGAGCGCCTCTGCGGCCGGCCCTGCCGCCCGGCCGCCGCCAC
GCCCAACCCGGCCCGCCCCCTAGCCCCGCCGGCCCGCCCCCGCCGCCGCCGCCAGGTGAGCGCTCCG
CCCGCCGAGGGCCCCGGCCCGCCCCCGCCCCCGCCGCCGCCGCCGCCAGGTGAGCGCTCCGCG
CGTCAAACCACTGATCCCATAAAACATTATCCTCCGGCCCGCCCTGCGCACCCGCCAGGCCAGTCCGCG
CGCCGCCGCCCTGCCCTGTGCGCCCTGCGCACCCGCCAGGCCAGCCAGAGCCGGCGGA
GCGGAGCGCCGAGCCTCGTCCGCCGGGGCGGGGCCGTAGCGCGCCCTGGATGCGGACCCG
GCCGCGGGAGACGGGCCCGCCCCGAAACGACTTTCAGTCCCAGCCGCCAACCCCTACGATGAA
GAGGGCGTCCGCTGGAGGGAGCCGGTGTGGCATGGGTGTGTGGCTGCAGGCCCTGGCAGGTGGCAGCCCAT
GCCCAAGGTGCTGCGTATGCTACAATGAGCCAAGGTGACGACAAGCTGCCAGCAGGGCTGCAGGCTGTG
CCCGTGGGATCCCTGCTGCCAGCGCATCTCCTGCACGGCACCGCATCTCGCATGTGCCAGCTGCCAG
CTTCCGTGCCCTGCCAACCTCACCATCCTGTGGCTGCACTCGAATGTGCTGGCCCGAATTGATGCGGCTGCC
TCACTGGCTGGCCCTCTGGAGCAGCTGGACCTCAGCGATAATGCACAGCTCCGGTGTGGACCCCTGCCACA
TTCCACGGCTGGGCCCTACACACGCTGCACCTGGACCCTGCAGGAGCTGGGGCCGGGGCTGTT
CCCGGGCTGGCTGCCCTGCAGTACCTCTACCTGCAGGACAACCGCTGCAGGACTGCCGTGATGACACCTCC
GCGACCTGGCAACCTCACACACCTCTTCTGCCAGGGCACCGCATCTCCAGCTGCCGGAGCGCCCTCCGT
GGGCTGCACAGCCTCGACCGTCTCCTACTGCACAGAAGCCGTTGGCCATGTGCACCCGATGCCCTGGCTGA
CCTTGGCCGCTCATGACACTCTATGTTGCAAAATCTATCAGCGCTGCCACTGAGGCCCTGGCCCTGGCCCC
TGGCTGCCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGGTGTGTGACTGCCGGCACGCCACTCTGGCC
TGGCTGAGAAGTTCGCCGGCTCCTCTCCGAGGTGCCCTGCAGCCTCCGCAACGCCCTGGCTGCCGTGACCT
CAAACGCCCTAGCTGCCAATGACCTGCAGGGCTGCGCTGTGGCCACCGGGCTTACCATCCATCTGGACCGCA
GGGCCACCGATGAGGAGCCGCTGGGCTCCAAAGTGCCTGCCAGCCAGATGCCGCTGACAAGGCCCTCAGTACTG
GAGCCTGGAAGACCAGCTCGGCAGGCAATGCGCTGAAGGGACGCGTGGCAGGGTGTGACAGCCGCCGGCAA
CGGCTCTGCCACGGCACATCAATGACTCACCTTGGACTCTGCCCTGGCTCTGCTGAGCCCCGGCTCACTG
CAGTGCAGGGCTCCGAGGCCACAGGGTCTCCACCTCGGGCCCTGCCGGAGGCCAGGTGTTCACGC
AAGAACCGCACCCGCAGCCACTGCCGTGGGCCAGGCAGGCCAGCAGGGGTGGCAGGACTGGTGAACAGAAGG
CTCAGGTGCCCTACCCAGCCTCACCTGCAGCCTACCCCCCTGGGCTGGCAGGTGTGACTGCCAGTGTGTT
GGCCCTGCTGAACCCAGGGACACAAGAGCGTGTGCTCAGCAGCCAGGTGTGACTACATACGGGGCTCTCTCCA
CGCCGCCAAGCCAGGCCGGGGCGGCCACCCGTGGGCAGGCCAGGTCTCCCTGATGGACGCCCTGCCGCC
CGCCACCCCATCTCCACCCATCATGTTACAGGGTTCGGCCGGCAGCCTTGTCCAGAACGCCCTCCAC
CCAGATCGGGTATATAGAGATATGCACTTTACTTGTGTAAAAATATCGGACGACGTGGAATAAGAGC
TCTTTCTTAAAAAA

130/246

FIGURE 128

MKRASAGGSRLLAWVLWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPVGIPAASQRIFLHGNRISHVPA
ASFRACRNLTILWLHSNVLARIDAAAFTGLALLEQLDLSDNAQLRSVDPATFHGLGLHHTLHLDRCGLQELGPG
LFRGLAALQYIYLQDNALQALPDDTFRDLGNLTHLFLHGNRISSVPERAFRGLHSLDRLLLHQNRVAHVPHAF
RDLGRLMTLYLFANNL SALPTEALAPLRLAQYLLNDNPWVCDCRARPLIWAWLQKFRGSSSEVPCSLPQRLAGR
DLKRLAANDLQGCAVATG PYHPIWTGRATDEEPPLGLPKCCQPDAA DKASVLEPGRPASAGNALKGRVPPGDSP
GNGSGPRHINDSPFGTLPGSAEPLTAVRPEGSEPPGFPTSGP RRRPGCSRKNRTRSHCRLGQAGSGGGGTGDS
EGSGALPSLTCSLTPLGLALVLWTVLGPC

Important features:

Signal peptide:

amino acids 1-26

Leucine zipper pattern.

amino acids 135-156

Glycosaminoglycan attachment site.

amino acids 436-439

N-glycosylation site.

amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

VWFC domain

amino acids 411-425

131/246

FIGURE 129

GGCGCCGGGAGCCCCATCTGCCCCCAGGGGCACGGGGCGCGGGGGCGGCTCCGCCCGGCACATGGCTGCAGCCAC
CTCGCGCGACCCCCGAGGCCCGCGCCAGCTCGCCCGAGGTCCCGTGGAGGCGCCGGCCGCCGGAGCCAA
GCAGCAACTGAGCGGGGAAGCGCCCGCGTCCGGGATGGGATGTCCTCCTCTCCTCTGCTAGTTCC
TACTATGTTGAAACCTGGGACTCACACTGAGATCAAGAGAGTGGCAGAGGAAAAGGTCACTTGCCCTGCC
CCATCAACTGGGCTTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTACCGATAATGAAGGGAACCAA
AAGTGGTGTACTTACTCCAGTCGTCTACATAACTTGACTGAGGAACAGAACGGCCAGTGGCTT
GCTTCAATTCTGGCAGGAGATGCCCTTGTGAGATTGAACCTCTGAAGGCCAGTGATGAGGCCGTACAC
CTGTAAGGTTAAGAATTCAAGGCGCTACGTGTGGAGCCATGTCATCTTAAAGTCTAGTGAGACCATCCAAGC
CCAAGTGTGAGTTGGAAGGGAGACTGACAGAAGGAAGTGACCTGACTTGTGAGTGATGAGTCACTCTGGCACA
GAGCCATTGTGATTACTGGCAGCGAATCGAGAGAAAGAGGGAGAGGATGAACGTCGTGCCCTCCAAATCTAG
GATTGACTACAACCACCCCTGGACGAGTCTGTCGAGAATCTTACCATGTCCTACTCTGACTGTACCGAGTC
CAGCAGGCAACGAAGCTGGAAAGGAAAGCTGTGTGGTGCAGTAACCTGACAGTATGTACAAAGCATCGGCATG
GTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGGCCCTGCTGATTTCTCTTGGGTGAGCTGCTAATCCGAAAG
GAAAGACAAAGAAAGATGAGGAAGAAGAGACCTAATGAAATTGAGAAGATGCTGAAGCTCCAAAGGCC
GTCTTGTGAAACCCAGCTCTTCCCTCAGGCTCTGGAGCTCAGCCTCTGGTCTCTCCACTCGCTCCACA
GCAAATAGTGCCTCACGCAGCCAGCGGACACTGTCACACTGACGCAGCACCCAGCCAGGGCTGGCCACCCAGGC
ATACAGCCTAGTGGGCCAGAGGTGAGAGTTCTGAACCAAAGAAAGTCCACCATGCTAATCTGACCAAAGCAG
AAACACACCCAGCATGATCCCCAGCCAGAGCAGGCCTCCAAACGGTCTGAATTACAATGGACTTACTCCC
ACGCTTCTAGGAGTCAGGGTCTTGGACTCTCTCGTCATTGGAGCTCAAGTCACCAGCCACACAACCAGAT
GAGAGGTCTAAGTAGCAGTGAGCATTGCAAGGAACAGATTAGCAGATGAGCATTCTTCTTACAATACCAAA
CAAGCAAAGGATGTAAGCTGATTCTGTGAGGACCTGTTGAGAAGGTTGGGAAAGGTGAGGTGAATATACCTAA
CAGGAGTCAAATCTATTGTTGACCAAGGACCTGTTGAGAAGGTTGGGAAAGGTGAGGTGAATATACCTAA
AACTTTAATGTGGATATTTGTATCAGTGTTGATTACAATTTCAAGAGGAATGGATGCTGTTGTA
AATTCTATGCAATTCTGCAAACCTATTGGATTATTAGTTATTGAGCTAAGACAGTCAGCAGAACCCACAGCCTTAT
TACACCTGCTACACCAGTACTGAGCTAACCAACTCTAAGAAACTCCAAAAAGGAAACATGTGCTTCTATT
CTGACTTAACCTCATTTGTCATAAGGTTGGATATTAAATTCAAGGGGAGTTGAAATAGTGGGAGATGGAGAAG
AGTGAATGAGTTCTCCACTCTATAACTAATCTACTATTGTTGAGGATGTTGACAAACATTAGAA
AAATTGTGACAAGGATTGTAAGAGCTTCCATCTCATGATGTTGAGGATGTTGACAAACATTAGAA
ATATATAATGGAGCAATTGTTGATTTCCCTCAATCAGATGCCCTAAGGACTTCTGCTAGATATTCTGG
AAGGAGAAAATACAACATGTCATTACACGTCCTTAGAAAGAATTCTCTAGAGAAAAGGGATCTAGGAAT
GCTGAAAGATTACCCAACATACCAATTAGTCTCTTCTGAGAAAATGTGAAACCAGAATTGCAAGACTGG
GTGGACTAGAAAGGGAGATTAGATCAGTTCTTAATATGTCAGGAAGGTAGCCGGCATGGTGCAG
CCTGAGGAAATCCAGCAGGTGGAGGTGCAGTGAGCCGAGATTATGCCATTGCACCCAGCTGGTGACAG
AGCGGGACTCCGTCTC

132/246

FIGURE 130

MSLLLLLLVSYYVGTTLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLTDNEGQKVVITYSSRHVYNN
LTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHVILKVLVRPSKPCELEGELTEGSD
LTIQCESSSGTIEPIVYYWQRIREKEGEDERLPPKSRIDYNHPGRVLLQNLTMYSGLYQCTAGNEAGKESCVVR
VTQYVQSIGMVAGAVTGIVAGALLIFLLVWLLIRRKDKEREEEERPNEIREDAEAPKARLVKPSSSSGSRS
SRGSSSTRSTANSASRSQRTLSTDAAAPQGLATQAYSLVGPEVRGSEPKVHHANLTKAETTPSMIPSQSRAFQTV

Important features:

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

133/246

FIGURE 131

GGAAGTCCACGGGAGCTGGATGCCAAGGGAGGACGGCTGGTCCCTGGAGAGGACTAC
TCACTGGCATATTCGAGGTATCTGTAGAATAACCACAGCCTCAGATACTGGGGACTTAC
AGTCCCACAGAACCGCCTCCCAGGAAGCTGAATCCAGCAAGAACAAATGGAGGCCAGCGGG
AGCTCATTGAGACAAAGGCAAGTCCTTTCTCCTTTGGGCTTATCTCTGGCG
GGCGCGGCGAACCTAGAAGCTATTCTGGTGGAGGAAACTGAGGGCAGCTCCCTTGTAC
CAATTAGCAAAGGACCTGGGCTGGAGCAGAGGAATTCTCCAGGCGGGGGTAGGGTTG
TTTCAGAGGGAAACAAACTACATTGAGCTCAATCAGGAGACCAGCGGATTGTTGCTAAAT
GAGAAATTGGACCGTGGAGGATCTGTGCGTCACACAGAGCCCTGTGTGCTACGTTCCAAGT
GTTGCTAGAGAGTCCCTCGAGTTTCAAGCTGAGCTGAAGTAATAGACATAAACGACC
ACTCTCCAGTATTCGGACAAACAAATGTTGGTGAAGAGTATCAGAGAGCAGTCCTCCTGG
ACTACGTTCTCTGAAGAATGCCAAGACTTAGATGTAGGCCAAACAAATATTGAGAACTA
TATAATCAGCCCCAACCTCTATTTCGGGTCCCTCACCGCAAACGCAAGTGTGAGCTGGCAGGAAAT
ACCCAGAGCTGGTGGACAAAGCGCTGGACCGAGAGGAAGAGCTGAGCTCAGGTTACA
CTCACAGCACTGGATGGTGGCTCCGCCAGATCTGGCACTGCTCAGGTCTACATCGAAGT
CCTGGATGTCAACGATAATGCCCTGAATTGAGCAGCCTTCTATAGAGTGCAGATCTCG
AGGACAGTCGGTAGGCTTCCTGGTTGAAGGTCTCTGCCACGGATGTAGACACAGGAGTC
AACGGAGAGATTCCATTCACTTTCAAGCTCAGAAGAGATGGCAAAACCTTAAGAT
CAATCCCTGACAGGAGAAATTGAACTAAAAAAACACTCGATTCGAAAAACTTCAGTCCT
ATGAAGTCAATATTGAGGCAAGAGATGCTGGAACCTTTCTGGAAAATGCACCGTTCTGATT
CAAGTGATAGATGTGAACGACCATGCCAGAAGTTACCATGTCTGCATTACAGCCAAAT
ACCTGAGAACGCGCCTGAAACTGTTGCACTTTCACTGTTTCACTGGTCTCAGATCTGATTAGGAG
AAAATGGGAAAATTAGTTGCTCCATTCAAGGAGGATCTACCCCTCCTGAAATCCGCGGAA
AACTTTACACCCACTAACGGAGAGACCACTAGACAGAGAAAGCAGAGCGGAATACAACAT
CACTATCACTGCACTGACTTGGGACCCCTATGCTGATAAACACAGCTCAATATGACCGTGC
TGATCGCCGATGTCAATGACAACGCTCCGCCCTCACCAAAACCTCTACACCCCTGTC
CGCGAGAACACAGCCCCGCCCTGCACATCCGCAAGCGTACAGACAGAGACTCAGG
CACCAACGCCAGGTACCTACTCGCTGCTGCCGCCCCAGGACCCGCACCTGCCCTCACAT
CCCTGGTCTCCATCAACGCGAACAGGCCACCTGTTGCCCTCAGGTCTCTGGACTACGAG
GCCCTGCAGGGGTTCCAGTTCCGCTGGCGCTCAGACCACGGCTCCCCGGCGCTGAGCAG
CGAGGCGCTGGTGCCTGGTGTGGTGTGGACGCCAACGACAACTCGCCCTCGTGTGTC
CGCTGCAGAACGGCTCCGCCCTGCACCGAGCTGGTCCCCGGCGGCCAGCCGGCTAC
CTGGTGACCAAGGTGGTGGCGGTGGACGGCGACTCGGGCAGAACGCCCTGGCTGTC
GCTGCTCAAGGCCACGGAGCTCGGTCTGTTGGCGTGTGGCGCACAATGGCGAGGTGCGCA
CCGCCAGGCTGCTGAGCGAGCGCAGCGGCCAACGACAGGCTGGTGGCTGGTCAAGGAC
AATGGCGAGCCTCCGCCCTGCCACGCCACGCTGCACTGCTCTGGTGGACGGCTTC
CCAGCCCTACCTGCCCTCCGGAGGCGCCACGCCAGGCCAGGGCAGCTGCTCACCG
TCTACCTGGTGGTGGCGTGGCCTCGGTGTCTCGCTCTCCCTCTTTCGGTGTGCTC
GTGGCGGTGCCCTGTTAGGAGGAGCAGGGCGGCCCTCGGTGGTCGCTGCTGGTGGCG
GGGCCCTCCAGGGCATCTGTGGACATGAGCGGCCACGCCAGGACCTATCCCAGAGCTACC
AGTATGAGGTGTGTCGGCAGGAGGCTCAGGGACCAATGAGTTCAAGTTCTGAAGCCGATT
ATCCCCAACCTCCCTCCCCAGTGCCCTGGAAAGAAATACAAGGAAATTCTACCTCCCCAA
TAACTTTGGGTTCAATATTCAGACCATAGTTGACTTTACATTCCATAGGTATTTATT
TGTGGCATTCCATGCCAATGTTATTCCCCAATTGTTGTTGTAATATTGTACGGAT
TTACTCTGATTTCATGTTCTCCCTTGTAAAGTGAACATTACCTTATT
CCTGGTTCTT

134/246

FIGURE 132

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48314
<subunit 1 of 1, 798 aa, 1 stop
<MW: 87552, pI: 4.84, NX(S/T): 5
MEASGKLICRQRQVLFSFLLGLSLAGAAEPRSYSVVEETEGSSFVTNLAKDLGLEQREFSR
RGVRVVSRGNKLHLQLNQETADLLLNEKLDREDLCGHTEPCVLRFQVILLESPPFEFFQAELOV
IDINDHSPVFLDKQMLVKVSESSPPGTTFPLKNAEDLDVGQNNIENYIISPNSYFRVLTRKR
SDGRKYPELVLDKALDREEEALRLTLTALDGGSPRSQTAQVYIEVLVDVNDNAPEFEQPFY
RVQISEDSPVGFLVVVKVSATDVDTGVNGEISYSLFQASEEIGKTFKINPLTGEIELKKQLDF
EKLQSYEVNIEARDAGTFSGKCTVLIQVIDVNDHAPEVTMSAFTSPIPENAPETVVALFSVS
DLDSENGKISCSIQEDLPFLLKSAENFYTLLTERPLDRESRAEYNITITVTDILGTPMLITQ
LNMTVLIADVNDNAPAFQTQSYTLFVRENNSPALHIRSVSATDRDSGTNAQVTYSLLPPQDP
HPLTLSLVSINADNGHFLALRSLDYEALQGFQFRVGASDHGSPALSSEALRVVVLDANDNS
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVVAVDGDSGQNAWLSYQLLKATELGLFGVWAH
NGEVRTARLLSERDAAKHRLVVLVKDNGEPPRSATATLHVLLVDGFSQPYLPLPEAAPTQAQ
ADLLTVYLVVALASVSSLFLFSVLLFVAVRLCRRSRAASVGRCLVPEGPLPGHLVDMSGTRT
LSQSYQYEVCLAGGSGTNEFKFLKPIIPNFPPQCPGKEIQGNSTFPNNFGFNIQ
```

Important features:**Signal peptide:**

amino acids 1-26

Transmembrane domain:

amino acids 685-712

Cadherins extracellular repeated domain signature.

amino acids 122-132, 231-241, 336-346, 439-449 and 549-559

ATP/GTP-binding site motif A (P-loop).

amino acids 285-292

N-glycosylation site.

amino acids 418-421, 436-439, 567-570 and 786-789

135/246

FIGURE 133

GGAAGGGAGGAGCAGGCCACACAGGCACAGGCCGGTGAGGGACCTGCCAGACCTGGAGGGTCTCGCTGTC
 ACACAGGCTGGAGTGCAGTGGTGTGATCTGGCTCATCGAACCTCCACCTCCGGTCAAGTGATTCTCATG
 CCTCAGCCTCCGAGTAGCTGGGATTACAGGTGGTGAATCCAAGAGTGACTCCGTCGGAGGAAAATGACTCCC
 CAGTCGCTGCTGCAGACACTGTTCTGCTGAGCTGCTCTCCGTCAAGGTGCCACGGCAGGGCCA
 CAGGGAGACTTCGCTCTGCAGCCAGCGGAACCAGACACACAGGAGCAGCCTCAACTACAAACCCACACAG
 ACCTGCGCATCTCCATCGAGAACCTCGGAAGAGGCCACAGTCCATGCCCTTCCCTGCAGGCCACCCGCT
 TCCGATCCTCCATCGAGGCCAGGGCCCTACACTCTGCTCTACTGGAACCGACATGCTGGGAGATTACA
 TCTCTCTATGGCAAGCGTGACTTCTGCTGAGTGACAAAGCCTTAGCCCTCTGCTCTCCAGCACCAGGAGG
 AGAGCCTGGCTCAGGGCCCCCGCTGTTAGCCACTCTGTCACCTCTGGTGGAGCCCTCAGAACATCAGCCTG
 CCCAGTGCCAGCTCACCTCTCCACAGTCCCTCCACAGGCCGCTCACATGCCCTGGTGGACAT
 GTGCGAGCTCAAAAGGGACCTCCAGCTGCTCAGCCAGTCTGAGACATCCCCAGAAGGCCCTAAGGAGGCC
 CGGCTGCCCGCCAGCCAGCAGTGCAGAGCTGGAGTCAGAAACTGACCTCTGTGAGATTATGGGGACATG
 GTGCTCTCGAGGAGGACGGATCAAGCCACGGTGTGGAGCTCCAGGCCACAGCCGGCTCAGGACCTGCA
 CATCCACTCCGGCAGGGAGGAGCAGAGCAGATCATGGAGTACTCGGTGCTGCTGCCGAACACTCTCC
 AGAGGACGAAAGGCCGGAGCAGGGCTGAGAAGAGACTCTCCTGGTGGACTTCAGCAGCCAAGGCCCTGTC
 CAGGACAAGAATTCCAGCCAAGTCTGGTGAGAAGGTCTTGGGATTGTGGTACAGAACACAAAGTAGCCAA
 CCTCAGGGAGCCCGTGGTGCTACTTCCAGCACCAGCTACAGCGAAGAATGTGACTCTGCAATGTGTTCT
 GGGTGAGACCCCCACATGAGCAGGCCGGGATTGGAGCAGTGTGCTGGTGTGAGACCGTCAGGAGAGAAACC
 CAAACATCTGCTCTGCAACCAACTTGACACTTGCAGTGTGCTGATGGTCTCTGGTGGAGGTGGACCGT
 GCACAAGCACTACCTGAGCCTCTCTACGGTGGCTGTGCTCTGCCCTGGCTGGCTGGTGTGACCATG
 CCGCCTACCTCTGCTCAGGGTGGCCATCTGCTGGACACGAGCTCTGCTCAGCGAGGCCGGTGGCCCTGACAGGCTCTGA
 AACCTGCTGCTGGCGCTTCTGCTGGACACGAGCTCTGCTCAGCGAGGCCGGTGGCCCTGACAGGCTCTGA
 GGCTGGCTGCCAGGCCATCTCTGCTCAGCTCTCTGCTCACCTGCCCTTCTGGATGGGCCCTGAGG
 GGTACAACCTCTACCGACTCGTGGTGGAGGTCTTGGCACCTATGTCCTGGTACACTACTCAAGCTGAGGCC
 ATGGGCTGGGCTTCCCATTTCTGGTACGCTGGTGGCCCTGGTGGATGTGGACAACATGGCCCATCAT
 CTTGGCTGTGCATAGGACTCCAGAGGGCGTACATCACCTCCATGTGCTGGATCCGGACTCCCTGGTCA
 ACATCACCAACCTGGGCCCTTCAGGCTGGTTCTGTTCAACATGGCATGCTAGCCACCATGGTGGTGCAG
 ATCCTGCGGCTGCGCCCCACACCCAAAAGTGGTACATGTGCTGACACTGCTGGGCCCTAGGCCGGTCTGG
 CCTGCCCTGGGCTTGATCTCTCTGGCACCTTCCAGCTTGTGCTCTGCCCTACCTTTCA
 TCATCACCTCTTCCAAGGCTTCTCATCTCATCTGGTACTGGTCCATGCCCTGAGGCCGGGGTGGCC
 TCCCTCTGAAGGCAACTCAGACAGGCCAGGCCCTCCATCAGCTGCCAGCACCTCGCCAGCCATTA
GGCCTCCAGCCCCACCTGGCCATGTGATGAAGCAGAGATGCCCTGGCAGACTGCCCTGAGGCCGGAGCC
AGGCCAGCCCCAGGCCAGTCAAGGCCAGACTTGGAAAGCCAACGACATGGAGAGATGGGCCCTGGCCATG
GTGGACGGACTCCGGCTGGCTTGAATTGGCCTTGGGACTACTCGGCTCTCAGCTCCCACGGGAC
TCAGAAGTGCAGGCCCATGCTGCCCTAGGTACTGTCCCCACATCTGTCCTCCACCCAGCTGGAGGCCGGTGTCT
CCTTACAACCCCTGGGCCAGGCCCTATTGCTGGGGCCAGGCCCTGGATCTGAGGGTCTGGCACATCTTAA
TCCGTGCCCTGCCCTGGGACAGAAATGTGGCTCAGTGTGCTGTCCTCTGTTGTCACCCCTGAGGGCACTCTG
CATCCTCTGTCATTTAACCTCAGGTGGCACCCAGGGCAATGGGCCAGGGCAGACCTCAGGCCAGAGCC
CTGGCGAGGAGAGGCCCTTGGCAGGAGCACAGCAGCAGCTGCCCTACCTCTGAGGCCAGGCCCTCC
CTCAGCCCCCAGTCCCTCCATCTCCCTGGGTTCTCTCTCTCCAGGCCCTCTGCTCCTCGTTC
ACAGCTGGGGTCCGATTCCAATGCTGTTTTGGGAGTGGTTCCAGGAGCTGCCCTGGTGTCTGCTGTA
ATGTTGCTACTGCACAAGCCTGCCCTGGCTGCCCCCTGAGCCAGGCTCGGTACCGATGCGTGGCTGGCTAGGTC
CCTCTGTCATCTGGGCTTGTATGAGCTGCATTGCCCTGCTCACCTGACCAAGCACAGGCCCTCAGAGGGG
CCCTCAGCCTCCCTGAAGGCCCTTGTGGCAAGAAACTGTGGACCATGCCAGTCCCCTGGTTCCATCCCAC
CACCTCCAAGGACTGAGACTGACCTCTCTGGTGAACACTGGCTAGGCCCTAGGCCCTGACACTCTCTCA
CAAGCCCCAAATAGCTCAGGCCCTGCCGCCCATCTGGTTAATTCTGTCACAAACACACAGGGTA
GATTGCTGGCTGTTGAGGTGGTAGGGACACAGATGACCGACCTGGTACTCTCTCTGCCAACATTCA
GTATGTGAGGGCTGCGTGAGCAAGAAACTCCTGGAGCTACAGGGACAGGGAGCCATATTCTGCC
CTGGAGACTCCCTGAGGAGTCAGCGTTCAATCTGACCTGAAAGATGGGAAGGGATGTTCTTACGTACCA
ATTCTTTGCTTTGATATTAAGAAGTACATGTTCATGTTAGAGAATTGGAAACTGTAGAAGAGAATCA
AGAAGAAAAAATAAAAATCAGCTGTTAATGCCCTAGCAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAA

136/246

FIGURE 134

MTPQSLLQTTLFLLSLLFLVQGAHGRGHREDFRFCSQRNQTHRSSLHYKPTPDLRISIENSE
EALTVHAPFPAAHPASRSFPDPRLGLYHFCLYWNRHAGRLHLLYGKRDFLSDKASSLLCFQH
QESLAQGPPLLATSVTSWSPQNISLPSAASFTFSFHSPPHTAAHNASVDMCELKRDLQLL
SQFLKHPQKASRRPSAAPASQQLQSLESKLTSVRFMGDMVSFEEDRINATVWKLQPTAGLQD
LHIHSRQEEEQSEIMEYSVLLPRTLFQRTKGRSGEAEKRLLLVDFSSQALFQDKNSSQVLGE
KVLGIVVQNTKVANLITEPVVLTQFHQLQPKNVTLCVFWVEDPTLSSPGHWSAGCETVRRE
TQTCFCNHLYFAVLMVSSVEVDQAVHKHYLSLLSYVGCVVSALACLVTIAAYLCSRVPILPC
RRKPRDYTIKVHMNLLLAVFLLDTSFLLSEPVALTGSEAGCRASAIFLHFSLTCLSWMGLE
GYNLYRLVVEVFGTYVPGYLLKLSAMGWGFPIFLVTIVALVDVDNYGPIILAVHRTPEGVIY
PSMCWIRDSDLVSYITNLGLFSLVFLFNAMLATMVVQILRLRPTHQKWSHVLTLLGLSLVLG
LPWALIFFSFASGTFQLVVLYLFSIITSFQGFLIFIWYWSMRLQARGGSPPLKSNSDSARLP
ISSGSTSSSRI

Important features:

Signal peptide:

amino acids 1-25

Putative transmembrane domains:

amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590
and 634-657

Microbodies C-terminal targeting signal.

amino acids 691-693

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 198-201 and 370-373

N-glycosylation sites.

amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327
and 341-344

G-protein coupled receptors family 2 proteins

amino acids 475-504

137/246

FIGURE 135

GCCTAGCCAGGCCAAGAATGCAATTGCCCGGTGGTGGGAGCTGGAGACCCCTGTGCTTGGACGGGACAGGGTCGG
GGGACACGCAGGATGAGGCCCGCGACCCTGGCACATTCTTGCTGACAGTGTACAGTATTTCTCCAAAGGTACA
CTCCGATCGGAATGTATAACCCATCAGCAGGTGCTCTTGTTCATGTTGGAAAGAGAATATTTAAGGGGG
AATTCCACCTTACCCAAAACCTGGCAGAGATTAGTAATGATCCCATAACATTAACTAAACAAATTAAATGGGTAC
CCAGACCGACCTGGATGGCTTCGATATATCCAAAGGACACCATATACTGATGGACTCCATATGGGTCCCCAAC
AGCTGAAAATGTGGGAAGCCAACAATCATTGAGATAACTGCCTACAAACAGGCGCACCTTGAGACTGCAAGGC
ATAATTGATAATTAATATAATGTCTGCAGAAGACTTCCCGTGCATATCAAGCAGAATTCTTCATTAAGAAT
ATGAATGTAGAAGAAATGTTGGCCAGTGGAGCTTGGAGACTTCTGGCGCAGTGGAAAATGTGTGGCAGCC
AGAGCGCCTGAACGCCATAAACATCACATCGGCCCTAGACAGGGGTGGCAGGGTGCACCTCCCCATTAAATGACC
TGAAGGAGGGCGTTATGTCATGGTGGTCAGATGTCCTTCTTCTTGTGAGAAGTGAAGGATCCA
CAGAATCAATTGAGATGTAGTCAAGAAATGGAGCCTGTAATAACATGTGATAAAAAATTCTGTACTCAATTAA
CATTGACTGGTGCAAATTCTGTTGATAAAACAAAGCAAGTGTCCACCTATCAGGAAGTGATTCTGTGGAG
AGGGGATTTACCTGATGGTGAGAATACAAACCCCTCTGATTCTTGAAAAGCAGAGACTATTACACGGAT
TTCTAATTACACTGGCTGTGCCCTCGGCAGTGCACTGGTCTTTCTAATACTGCTTATATCATGTGCTG
CCGACGGGAAGGCCTGGAAAAGAGAAACATGCAAACACCAAGACATCCAACGGTCCATCACAGTGTCTTCA
AATCTACCAAGGAGCTCGAGACATGTCCAAGATAAGAGATAGCATGGCCCTGTCAACGCTCCTGTGTT
CACCTGTGACTGGGAAATCATACCTCTTACACAGACAACATGATAGCACAACATGCCATTGATGCA
AACGCAGCAGAACTTGCCACATCAGACTCAGATCCCCAACAGCAGACTACAGGTAAATGGTATCCCTGAAGAA
AGAAAATGACTGAAGCAATGAATTATAATCAGACAATATGAGCTACATCACATTCTTTCTCTTCAAT
AATGCATGAGCTTTCTGGCATATGTTATGCTGAGTATTAGTGATACCAAAATAATACAACATAACT
TTCATTAACTAATGTTTTTGACTTAAAGCATTGGACAATTGAAACATTGATGACTTATATT
GTTACAATAAAAGTTGATCTTAAATAAATATTATTAATGAAGCCTAAAAAA

138/246

FIGURE 136

MQLPRWWELGDPCAWTGQGRGTRRMSATTGTFLLTVYSIFSKVHSDRNVYPSAGVLFVHVLEREYFKGEFPPY
PKPGEISNDPITFNTNLMGYPDRPGWLRYIQRTPYSDGVLYGSPTAENVGKPTIIIEITAYNRRTFETARHNLII
NIMSAEDFPLPYQAEFFIKNMNVEEMLASEVLGDFLGAVKNNWQPERLNAINITSALDRGGRVPLPINDLKEGV
YVMVGADVPFSSCLREVENPQNQLRCSQEMEPVITCDKKFRQFYIDWCKISLVDKTKQVSTYQEVRGEGILP
DGGEYKPPSDSLKSRDYYTDFLITLAVPSAVALVLFLILAYIMCCRREGVEKRNMQTPDIQLVHHSAIQKSTKE
LRDMSKNREIAWPLSTLPVFHPVTGEIIPPLHTDNYDSTNMPLMQTQQNLPHQTQIPQQQTTGKWYP

signal sequence:
Amino acids 1-46

transmembrane domain:
Amino acids 319-338

N-glycosylation site:
Amino acids 200-204

cAMP- and cGMP-dependent protein kinase phosphorylation site:
Amino acids 23-27

Tyrosine kinase phosphorylation site:
Amino acids 43-52

N-myristoylation sites:
Amino acids 17-23; 112-118; 116-122;
185-191

139/246

FIGURE 137

CAGAAGAGGGGGCTAGCTAGCTCTGCGGACCAGGGAGACCCCCCGCGCCCCCGGTGT
GAGGCGGCCACAGGGCGGGTGGCTGGCAGCCGACGCCGGCGAGGAGGCTGTGAG
GAGTGTGTGGAACAGGGACCCGGGACAGAGGAACCATGCTCCGCAGAACCTGAGCACCTTT
GCCTGTTGCTGCTATAACCTCATCGGGCGGTGATTGCCGGACGAGATTTCTATAAGATCTG
GGGTGCCTCGAAGTGCCTCTATAAAGGATATTAAAAGGCCTATAGGAAACTAGCCCTGCA
GCTTCATCCCACCGAACCCCTGATGATCCACAAGCCCAGGAGAAATTCCAGGATCTGGGTG
CTGCTTATGAGGTTCTGTCAAGATAGTGAGAACAGTACGATACTTATGGTGAAGAA
GGATTAAGATGGTCATCAGAGCTCCCATGGAGACATTTTCAACTCTTGGGATT
TGGTTCATGTTGGAGGAACCCCTCGTCAGCAAGACAGAAATATTCAAGAGGAAGTGATA
TTATTGTAGATCTAGAAGTCACTTGGAGAAGTATATGCAGGAAATTGTGAAGTAGTT
AGAAACAAACCTGTGCAAGGCAGGCTCTGGCAACCGGAAGTGCAATTGTCGGCAAGAGAT
GCGGACCACCCAGCTGGGCCCTGGCGCTTCCAATGACCCAGGAGGGTCTGCGACGAAT
GCCCTAATGTCAAACTAGTGAATGAAGAACGAACCGCTGGAAGTAGAAATAGAGCCTGGGTG
AGAGACGGCATGGAGTACCCCTTATTGGAGAAGGTGAGCCTCACGTGGATGGGAGCCTGG
AGATTTACGGTTCCGAATCAAAGTTGTCAAGCACCAATATTGAAAGGAGAGGAGATGATT
TGTACACAAATGTGACAATCTCATTAGTGAGTCAGTGGCTGGCTTGAGATGGATATTACT
CACTGGATGGTCACAAGGTACATATTCGGGATAAGATCACCAGGCCAGGAGCGAAGCT
ATGGAAGAAAGGGGAAGGGCTCCCAACTTGACAAACAATATCAAGGGCTTTGATAA
TCACTTTGATGTGGATTTCAAAAGAACAGTTAACAGAGGAAGCGAGAGAAGGTATCAA
CAGCTACTGAAACAAGGGTCAGTGCAGAAGGTATAACATGGACTGCAAGGATATTGAAGTG
AATAAAATTGGACTTGTAAATAAGTGAATAAGCGATATTATTATCTGCAAGGTTTT
TTGTGTGTGTTTTGTTTATTTCATATGCAAGTTAGGCTTAATTGTTATCTAATGA
TCATCATGAAATGAATAAGAGGGCTTAAGAATTGTCCATTGCATTGGAAAAGAATGACC
AGCAAAAGGTTACTAATACCTCTCCCTTGGGATTTAATGTCTGGCTGCCCTGAGT
TTCAAGAATTAAAGCTGCAAGAGGACTCCAGGAGCAAAAGAAACACAATATAGAGGGTTGGA
GTTGTTAGCAATTCAAAATGCCAATGGAGAAGTCTGTTTAAATACATTTGTT
TTATTTTA

140/246

FIGURE 138

MAPQNLSTFCLLLLYLIGAVIAGRDFYKILGVPRSASIKDIKKAYRKLAQLHPDRNPDDPQAQEKFQDLGAAY
EVLSDSEKRKQYDTYGEGLKDGHQSSHGDIFSHFFGDFGFMFGGTPROQDRNIPRGSDIIVDILEVTLEEVYAG
NFVEVVRNKPVARQAPGKRKCNCRQEMRTTQLGPGRFQMTQEVVCDCECBNVKLVNEERTLEVEIEPGVRDGMEY
PFIGECEPHVDGEPEGDILRFRIKVVKHPIFERRGDDLYTNVTISLVESLVGFEDEMITHLDGHKVHISRDKITRPG
AKLWKKGEGLPNFDNNNIKGSLIITFDVDFPKEQLTEEAREGIKQLLKQGSVQKVYNGLQGY

Important features:

Signal peptide:

amino acids 1-22

Cell attachment sequence.

amino acids 254-257

Nt-dnaJ domain signature.

amino acids 67-87

Homologous region to Nt-dnaJ domain proteins.

amino acids 26-58

N-glycosylation site.

amino acids 5-9, 261-265

Tyrosine kinase phosphorylation site.

amino acids 253-260

N-myristoylation site.

amino acids 18-24, 31-37, 93-99, 215-221

Amidation site.

amino acids 164-168

141/246

FIGURE 139

CCAGTCTGTCGCCACCTCACTGGTGTCTGCTGTCCCCGCCAGGCAAGCCTGGGTGAGAGC
ACAGAGGAGTGGGCCGGGACCATGCGGGGGACGCCGCTGGCGCTCTGGCGCTGGTGC
TGCCTGCGGAGAGCTGGCGCCCTGCGCTGCTACGTCTGTCCGGAGCCCACAGGAGTGT
CGGACTGTGTCACCATGCCACCTGCACCAACGAAACCATGTGCAAGACCACACTCTAC
TCCCAGGGAGATAGTGTACCCCTTCCAGGGGACTCCACGGTGACCAAGTCCCTGTGCCAGCAA
GTGTAAGCCCTCGGATGTGGATGGCATCGGCCAGACCCCTGCCGTGCTGCAATACTG
AGCTGTGCAATGTAGACGGGCGCCCGCTCTGAACAGCCTCCACTGCCGGGCCCTACGCTC
CTCCCACTCTTGAGCCTCCGACTGTAGAGTCCCCGCCACCCCATGCCCTATGCCGGCCA
GCCCGAATGCCCTGAAGAAGTGCCCTGCACCAAGGAAAAAAAAAAAAAAA

142/246

FIGURE 140

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56405
<subunit 1 of 1, 125 aa, 1 stop
<MW: 13115, pI: 5.90, NX(S/T): 1
MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTTLYSREIVYP
FQGDSTVTKSCASKCKPSDVDGIGQTLPVSCCNTELNVGDGAPALNSLHCGALTLLPLSLRL
```

Important features:

Signal peptide:

amino acids 1-17

N-glycosylation site.

amino acids 46-49

143/246

FIGURE 141

GGCGCCGCGTAGGCCCGGGAGGCCGGCCGGCTGCGAGCGCCTGCCCATGCGCCGC
CGCCTCTCCGCACGATGTCCCCCTCGCGAGGAAAGCGCGCAGCTGCCCTGGGAGGACGGC
AGGTCCGGGTTGCTCTCGCGGGCCTCCCTCGGAAGTGTCCGTCTCCACCTGTTCGTGGC
CTGCCTCTCGCTGGGCTTCTTCTCCCTACTCTGGCTGCAGCTCAGCTGCTCTGGGAGCTGG
CCCAGGCAGTCAGGGGACAAGGGCAGGAGACCTCGGGCCCTCCCCGTGCCTGCCCCCCAGAG
CCGCCCCCTGAGCAGTGGAGAAAGACGCATCCTGGGCCCCCACCGCCTGGCAGTGCTGGT
GCCCTCCCGAACGCTCGAGGAGCTCTGGTCTCGTCCCCACATGCGCCGCTTCTGA
GCAGGAAGAAGATCCGGCACACATCTACGTGCTCAACCAGGTGGACCACTTCAGGTTAAC
CGGGCAGCGCTCATCAACGTGGCTTCCCTGGAGAGCAGCAACAGCACGGACTACATTGCCAT
GCACGACGTTGACCTGCTCCCTCTCAACGAGGAGCTGGACTATGGCTTCCCTGAGGCTGGC
CCTTCCACGTGGCTCCCCGGAGCTCCACCCCTCTACCAACTACAAGACCTATGTCGGCGGC
ATCCTGCTGCTCTCCAAGCAGCACTACCGGCTGTGCAATGGGATGTCCAACCGCTTCTGGGG
CTGGGGCCCGAGGACGAGTCTACCGGCGCATTAAGGGAGCTGGCTCCAGCTTTCC
GCCCTCGGGAAATACAACACTGGTACAAGACATTGCCACCTGCATGACCCAGCCTGGCGG
AAGAGGGACCAAGCGCATCGCAGCTAAAAACAGGAGCAGTTCAAGGTGGACAGGGAGGG
AGGCCTGAACACTGTGAAGTACCATGTGGCTCCCGACTGCCCTGTCTGTGGGCGGGCCC
CCTGCACTGTCCTCAACATCATGTTGGACTGTGACAAGACGCCACACCCCTGGTCACATTC
AGCTGAGCTGGATGGACAGTGAGGAAGCCTGTACCTACAGGCCATATTGCTCAGGCTCAGGA
CAAGGCCCTAGGTGTTGGCCAGCTCTGACAGGATGTGGAGTGGCCAGGACCAAGACAGCA
AGCTACGCAATTGCAGCCACCCGGCCGCAAGGCAGGCTTGGCTGGGCCAGGACACGTGGG
GTGCCTGGGACGCTGCTGCCATGCACAGTGTGATCAGAGAGAGGGCTGGGTGTGTCCTGTCCG
GGACCCCCCTGCCCTCCCTGCTCACCTACTCTGACCTCCTCACGTGCCAGGCCGTGG
TAGTGGGGAGGGCTGAACAGGACAACCTCTCATCACCTACTCTGACCTCCTCACGTGCC
AGGCCTGTGGTAGTGGGAGGGCTGAACAGGACAACCTCTCATCACCCCCAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

144/246

FIGURE 142

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56531
><subunit 1 of 1, 327 aa, 1 stop
><MW: 37406, pI: 9.30, NX(S/T): 1
MFPSRRKAAQLPWEDGRSGLLSGGLPRKCSVFHLFVACLSLGFFSLLWLQLSCSGDVARAVR
GQQQETSGPPRACPPPEHWEEDASWGPRLAVLVPFRERFEELLVFVPHMRRFLSRKKI
RHIYVLNQVDHFRFNRAALINVGFLESSNSTDYIAMHDV DLLPLNEELDYGFPEAGPFHVA
SPELHPLYHYKTYVGGILLLSKQHYRLCNGMSNRFWGWGREDDEFYRIKGAGLQLFRPSGI
TTGYKTFRHLHDPAWRKRDQKRIAAQKQEKFVDRREGGLNTVKYHVASRTALSVGGAPCTVL
NIMLDCKTATPWCTFS
```

Signal peptide:

amino acids 1-42

Transmembrane domain:

amino acids 29-49 (type II)

N-glycosylation site.

amino acids 154-158

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 27-31

Tyrosine kinase phosphorylation site.

amino acids 226-233

N-myristoylation site.

amino acids 19-25, 65-71, 247-253, 285-291, 303-309, 304-310

145/246

FIGURE 143

GTGGGATTATTGAGTGCAGATCGTTCTCAGTGGGGAGTTGCCTCATCGCAGGCAGATGTTGGGG
CTTGTCCGAACAGCTCCCTCTGCCAGCTCTGTAGATAAGGGTAAAAAACTAATATTATGACAGAAGAA
AAAGATGTCATTCCGTAAAGTAAACATCATCATCTGGTCTGGCTGTGCTCTCTTACTGGTTTGACCC
ATAACTCCCTCAGCTTGAGCAGTTGTTAAGGAATGAGGTACAGATTAGGAATTGTTAGGGCCTAACCTATA
GACTTGTCCAAATGCTCTCCGACATGCAGTAGATGGGAGACAAGAGGAGATTCTGTGGTCATCGCTGCATC
TGAAGACAGGGTTGGGGGCCATTGCAGCTATAAACAGCATTAGCACAACACTCGCTCCAATGTGATTTCT
ACATTGTTACTCTCAACAATACAGCAGACCATTCCGGTCTGGCTCAACAGTGATTCCCTGAAAAGCATCAGA
TACAAAATTGCAATTGGACCCCTAAACTTGGAAAGGAAAGTAAAGGAGGATCTGACCAGGGGAATCCAT
GAAACCTTAACCTTGCAAGGTTCTACTTGCAATTCTGGTCCCAGCGCAAAGAAGGCCATATACATGGATG
ATGATGTAATTGTGCAAGGTGATATTCTGCCCTTACAATACAGCACTGAAGCCAGGACATGCAGCTGCATT
TCAGAAGATTGTGATTGCCTACTAAAGTTGTCATCCGTGGAGCAGGAAACCACTACAATTACATTGGCTA
TCTGACTATAAAAGGAAAGAATTGTAAGCTTCCATGAAAGCCAGCACTTGCTCATTTAATCCCTGGAGTTT
TTGTTGCAAACCTGACGGAATGGAACGACAGAAATATAACTAACCAACTGGAAAATGGATGAAACTCAATGTA
GAAGAGGGACTGTATAGCAGAACCTGGCTGGTAGCATCACAAACACCTCCTCTGCTTATCGTATTTATCAACA
GCACCTACCATCGATCCTATGGAAATGTCGCCACCTGGTCCAGTGCTGGAAAACGATATTACCTCAGT
TTGTAAGGCTGCCAGTTACTCCATTGGAAATGGACATTGAAAGCCATGGGAAGGACTGCTTCATATACTGAT
GTTGGGAAAATGGTATATTCCAGACCAACAGGCAAATTCAACCTAATCGAAGATATACCGAGATCTCAA
CATAAAGTGAAACAGAATTGAACTGTAAGCAAGCATTCTCAGGAAGTCCTGGAAGATAGCATGCATGGGAAG
TAACAGTTGCTAGGCTCAATGCCTATCGGTAGCAAGCCATGGAAAAGATGTGTCAGTAGGTAAGATGACA
AACTGCCCTGTCGGCAGTCAGCTCCAGACAGACTATAGACTATAAAATATGTCCTCATGCTCTACCAAGT
GTTTCTTACTACAATGCTGAATGACTGGAAAGAAGACTGATATGGCTAGTTGAGCTAGCTGGTACAGATAAT
TCAAAACTGCTGTTGGTTAATTGTAACCTGTCGATCTGTAATAAAACTACATTTTC

146/246

FIGURE 144

MSFRKVNIILVLVALFLLVLHHNFLSLLRNEVTDSGIVGPQPIDFVPNALRHADGR
QEEIPVVIASEDRLLGGAIAAINSIQHNTRSNVIFYIVTLNNTADHLRSWLNSDSLKSIRYK
IVNFDPKLLEGKVKEDPDQGESMKPLTFARFYLPILVPSAKKAIYMDDDVIVQGDILALYNT
ALKPGHAAAFSEDCDSASTKVVIRGAGNQNYIGYLDYKKERIRKLSMKASTCSFNPGVFVA
NLTEWKRNQITNQLEKWMKLNVEEGLYSRTLAGSITTPPLLIVFYQQHSTIDPMWNVRHLGS
SAGKRYSPQFVKAAKLLHWNGHLKPWGRTASYTDVWEKWYIPDPTGKFNLIRRTEISNIK

147/246

FIGURE 145

AAACTTGACGCCATGAAGATCCC GGTCCTT CCTGCCGTGGTGCTCCTCTCCCTCCTGGTGCT
CCACTCTGCCAGGGAGCCACCCCTGGGTGGTCTGAGGAAGAACCCATTGAGAATTATG
CGTCACGACCCGAGGCCTTAAACACCCGTTCTGAACATCGACAAATTGCGATCTGCGTT
AAGGCTGATGAGTTCTGAAC TGGCACGCCCTTTGAGTCTATCAAAGGAAACTTCCTT
CCTCAACTGGGATGCCTTCCTAAGCTGAAAGGACTGAGGAGCGCAACTCCTGATGCCAGT
GACCATGACCTCCACTGGAAGAGGGGGCTAGCGTGAGCGCTGATTCTAACCTACCATAACT
CTTCCTGCCTCAGGAACCTCAAATAAACATTTCATCCAAA

148/246

FIGURE 146

MKIPVLPAVVLLSLLVLHSAQGATLGGPEEESTIENYASRPEAFNTPFLNIDKLRSAFKADE
FLNWHALESIKRKLPFLNWDAFPKLKGLRSATPDAQ

149/246

FIGURE 147

CCTCTGTCCACTGCTTCGTGAAGACAAGATGAAGTTCACATTGTCTTGCTGGACTTCTT
GGAGTCTTCTAGCTCCTGCCCTAGCTAACTATAATCAACGTCAATGATGACAACAACAA
TGCTGGAAGTGGGCAGCAGTCAGTGAGTGTCAACAATGAACACAATGTGGCCAATGTTGACA
ATAACAACGGATGGGACTCCTGGAATTCCATCTGGGATTATGGAATGGCTTGCTGCAACC
AGACTCTTCAAAAGAAGACATGCATTGTCACAAAATGAACAAGGAAGTCATGCCCTCCAT
TCAATCCCTGATGCACTGGTCAAGGAAAAGAACGTTCAGGGTAAGGGACCAGGAGGACCAC
CTCCAAGGGCCTGATGTACTCAGTCACCCAAACAAAGTCGATGACCTGAGCAAGTCGGA
AAAAACATTGCAAACATGTGTCGTGGATTCCAACATACATGGCTGAGGAGATGCAAGAGGC
AAGCCTGTTTTTACTCAGGAACGTGCTACACGACCAGTGTACTATGGATTGGACATT
CCTCTGTGGAGACACGGTGGAGAAACTAAACAATTAAAGCCACTATGGATTAGTCAT
CTGAATATGCTGTGCAGAAAAATATGGCTCCAGTGGTTTACCATGTCATTCTGAAATT
TTTCTCTACTAGTTATGTTGATTCTTAAGTTCAATAAAATCATTAGCATTGAAAAAA

150/246

FIGURE 148

MKFTIVFAGLLGVFLAPALANYNINVNDNNNAGSGQQSVNNEHNVANVDNNNGWDSWNS
IWGYNGFAATRLFQKKTCIVHKMNEVMPSIQSLDALVKEKKLQGKPGGPPPGLMYSVN
PNKVDDLSKFGKNIANMCRGIPTYMAEEMQEASLFFYSGTCYTSVLWIVDISFCGDTVEN

Signal Peptide:

amino acids 1-20

N-myristoylation Sites:

amino acids 67-72, 118-123, 163-168

Flavodoxin protein homology:

amino acids 156-174

151/246

FIGURE 149

GGCACGAGCCAGGAACCTAGGAGGTTCTCACTGCCGAGCAGAGGCCCTACACCCACCGAGGC
ATGGGCTCCCTGGCTGTTCTGCTTGGCGTGCTGGCTGCCAGCAGCTTCTCCAAGGCACG
GGAGGAAGAAATTACCCCTGTGGTCTCATTGCCACAAAGCTGGAAGTTTCCCCAAAG
GCCGCTGGGTGCTCATAACCTGCTGTGCACCCCAGCCACCACCGCCCACCTACCTATTCCCTC
TGTGGAACCAAGAACATCAAGGTGGCAAGAAGGTGGTAAGACCCACGAGCCGGCCTCCTT
CAACCTCAACGTCACACTCAAGTCCAGTCCAGACCTGCTCACCTACTTCTGCCGGGCGTCCT
CCACCTCAGGTGCCCATGTGGACAGTGCAGGCTACAGATGCACGTGGAGCTGTGGTCCAAG
CCAGTGTCTGAGCTGCCGGCCAACTTCACTCTGCAGGACAGAGGGCAGGCCAGGGTGG
GATGATCTGCCAGGGCTCCTCGGGCAGCCCACCTATCACCAACAGCCTGATCGGGAGGATG
GCCAGGTCCACCTGCAGCAGAGACCATGCCACAGGCAGCCTGCCAACCTCTCCTCGCG
AGCCAGACATCGGACTGGTTCTGGTGCCAGGCTGCAAACAAAGCCAATGTCCAGCACAGCG
CCTCACAGTGGTGCCCGAGGTGGTGACCAAGATGGAGGACTGGCAGGGTCCCCTGGAGA
GCCCATCCTGCCCTGCCGCTCTACAGGAGCACCGCCGTCTGAGTGAAGAGGGAGTTGGG
GGTTCAAGGATAGGGAAATGGGAGGTCAAGGGACGCAAAGCAGCAGCCATG**TAGA**ATGAACC
GTCCAGAGAGCCAAGCACGGCAGAGGACTGCAGGCCATCAGCGTGCACGTGTTGTATTTGGA
GTTCATGCAAAATGAGTGTGTTAGCTGCCACAAAAAAAAAAAAAAA

152/246

FIGURE 150

MGLPGLFCLAVLAASSFSKAREEEITPVVSIAYKVLEVFPKGRWVLITCCAPQPPPITYSL
CGTKNIKVAKKVVKTHEPASFNLNVTLKSSPDLLTYFCRASSTSGAHVDSARLQMHWELWSK
PVSELRANFTLQDRGAGPRVEMICQASSGSPPITNSLIGKDGQVHLQQRPCHRQPANFSFLP
SQTSDFWFCQAANNANVQHSALTVVPPGGDQKMEDWQGPLESPILALPLYRSTRRLSEEEFG
GFRIGNGEVGRKAAAM

Signal Peptide:
amino acids 1-18

N-glycosylation Sites:
amino acids 86-89, 132-135, 181-184

153/246

FIGURE 151

GCGTGGGG**ATGT**CTAGGAGCTCGAAGGTGGTCTGGGCCTCTCGGTGCTGCTGACGGCGGCC
ACAGTGGCCGGCGTACATGTGAAGCAGCAGTGGGACCAGCAGAGGCTTCGTGACGGAGTTAT
CAGAGACATTGAGAGGCAAATTGGAAAAAAAGAAAACATTCGTCTTTGGGAGAACAGATTA
TTTGACTGAGCAACTTGAAGCAGAAAGAGAGAAGATGTTATTGGCAAAAGGATCTCAAAA
TCAT**TGA**CTTGAATGTGAAATATCTGTTGGACAGACAAACACGAGTTGTGTGTGTTGAT
GGAGAGTAGCTTAGTAGTATCTTCATCTTTTTGGTCACTGTCCTTTAAACTTGATCA
AATAAAGGACAGTGGGTCAATAAGTTACTGCTTCAGGGTCCCTTATATCTGAATAAAGGA
GTGTGGGCAGACACTTTGGAAGAGTCGTCTGGGTGATCCTGGTAGAAGCCCCATTAGGG
TCACTGTCCAGTGCTTAGGGTTACTGAGAAGCACTGCCAGCTTGTGAGAAGGAAGGG
TGGATAGTAGCATCCACCTGAGTAGTCTGATCAGTCGGCATGATGACGAAGCCACGAGAAC
TCGACCTCAGAAGGACTGGAGGAAGGTGAAGTGGAGGGAGAGACGCTCTGATCGTCGAATCC

154/246

FIGURE 152

MSRSSKVVLGLSVLLTAATVAGVHVQQWDQQQLRDGVIRDIERQIRKKENIRLLGEQIILT
EQLEAEREKMLLAKGGSQKS

155/246

FIGURE 153

AATGTGAGAGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTAGCACCAGTACTGGAT
GTGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTCACTGTCCGATTCTGATTCCGGCAAGG
ATCCAAGCATGGAATGCTGCCGTGGCAACTCCTGGCACACTGCTCTCTTCTGGCTTC
CTGCTCCTGAGTTCAGGACCGCACGCTCCGAGGAGGACGGGACGGCTATGGGATGCCCTG
GGGCCATGGAGTGAATGCTCACGCACCTGCGGGGAGGGCCTACTCTGAGGCCT
GCCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATAACAGAACATGCAGTAATGTGGAC
TGCCCACCAAGCAGGTGATTCCGAGCTCAGCAATGCTCAGCTATAATGATGTCAAGCA
CCATGCCAGTTTATGAATGCCCTGTCTAATGACCCCTGACAACCCATGTTACTCA
AGTGCCAAGCCAAAGGAACAACCCCTGGTTGTGAACTAGCACCTAAGGTCTTAGATGGTACG
CGTTGCTATACAGAATCTTGGATATGTGCATCAGTGGTTATGCCAAATTGTTGGCTGCGA
TCACCAGCTGGGAAGCACCGTCAGGAAGATAACTGTGGGTCTGCAACGGAGATGGTCCA
CCTGCCGGCTGGTCCGGAGGGCAGTATAAATCCCAGCTCTCCGCAACCAAATGGATGATACT
GTGGTTGCACTCCCTATGGAAGTAGACATATTGCCCTGTCTTAAAGGTCTGATCACTT
ATATCTGAAACCAAAACCCCTCAGGGACTAAAGGTGAAAACAGTCAGCTCCACAGGAA
CTTCCTTGTGGACAATTCTAGTGTGGACTTCCAGAAATTCCAGACAAAGAGATACTGAGA
ATGGCTGGACCACTCACAGCAGATTTCATTGTCAAGATTGTAACCTGGGCTCCGCTGACAG
TACAGTCCAGTTCATCTTCTATCAACCCATCATCCACCGATGGAGGGAGACGGATTCTTC
CTTGCTCAGCAACCTGTGGAGGGTTATCAGCTGACATCGGCTGAGTGCTACGATCTGAGG
AGCAACCGTGTGGTTGCTGACCAATACTGTCACTATTACCCAGAGAACATCAAACCCAAACC
CAAGCTTCAGGAGTGCAACTTGGATCCTGTCCAGCCAGTGACGGATAACAAGCAGATCATGC
CTTATGACCTCTACCACATCCCCTCCTCGGTGGAGGCCACCCATGGACCGCGTCTCC
TCGTGTGGGGGGGCATCCAGAGCCGGCAGTTCTGTGTGGAGGGACATCCAGGGCA
TGTCACTTCAGTGGAAAGAGTGGAAATGCATGTACACCCCTAAAGATGCCATCGCGCAGCCCT
GCAACATTTTGACTGCCCTAAATGGCTGGCACAGGAGTGGCTCCGTGCACAGTGACATGT
GCCAGGGCCTCAGATACCGTGTGGCTCTGCATCGACCATCGAGGAATGCACACAGGAGG
CTGTAGCCCCAAAACAAAGCCCCACATAAAAGAGGAATGCATGTAACCACTCCCTGCTATA
AACCCAAAGAGAAACTCCAGTCGAGGCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTA
GAAGAAGGAGCTGCTGTCAAGAGGCCCTCGTAAGTTGAAAAGCACAGACTGTTCTATA
TTGAAACTGTTTGTAAAGAAAGCAGTGTCACTGGTTGTAGCTTCAATTAAAGATTGATTGTTCAA
AAAAAA

156/246

FIGURE 154

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58847
<subunit 1 of 1, 525 aa, 1 stop
<MW: 58416, pI: 6.62, NX(S/T): 1
MECCRRATPGTLLLFLAFLSSRTARSEEDRDGLWDAGPWSECSRTCGGGASYSLRRCLS
SKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVHHGQFYEWLPVSNDPDNPCSLKCQ
AKGTTLVVELAPKVLGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCGDGSTCR
LVRGQYKSQLSATKSDDTVVALPYGSRHIRLVLKGPDHLYLETKTLQGTKGENSLSSGTFL
VDNSSVDFQKFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFPCS
ATCGGGYQLTSAECYDLRSNRVADQYCHYPENIKPKPKLQECNLDPASPASDGYKQIMPYD
LYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNI
FDCPKWLAQEWSPTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPK
EKLPEAKLPWFQQAQEEGAAVSEEPS
```

Important features:**Signal peptide:**

amino acids 1-25

N-glycosylation site.

amino acids 251-254

Thrombospondin 1

amino acids 385-399

von Willebrand factor type C domain proteins

amino acids 385-399, 445-459 and 42-56

FIGURE 155

158/246

FIGURE 156

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59212
><subunit 1 of 1, 440 aa, 1 stop
><MW: 42208, pI: 6.36, NX(S/T): 1
MKFQGPLACLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGA
AGSKVSEALGQGTREAVGTGVRQVPGFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHG
ADAVRGSWQGVPGHSGAWETSGGHGIFGSQGGLGGQGQGNPGGLGTPWVHYPGNSAGSFGM
NPQGAPWGQGGNGGPPNFGTNTQGAVAQPGYGSVRASNQNEGCTNPPPSGSGGSSNSGGGS
GSQSGSSGSGSNGDNNNNGSSSGSSSGSSGGSSGGSSGSSGNGSGSRGDSGSESSW
GSSTGSSSGNHGGSGGNHKGCEKPGNEARGSGESGIQGFRGQGVSSNMREISKEGNLL
GSSGDNYRGQGSSWGSGGGDAVGGVNTVNSETSPGMFNFDTFWKNFKSKLGFINWDAINKDQ
RSSRIP
```

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70,
74-80, 90-96, 96-102, 130-136, 140-146, 149-155, 152-158,
155-161, 159-165, 163-169, 178-184, 190-196, 194-200, 199-205,
218-224, 236-242, 238-244, 239-245, 240-246, 245-251, 246-252,
249-252, 253-259, 256-262, 266-272, 270-276, 271-277, 275-281,
279-285, 283-289, 284-290, 287-293, 288-294, 291-297, 292-298,
295-301, 298-304, 305-311, 311-317, 315-321, 319-325, 322-328,
323-329, 325-331, 343-349, 354-360, 356-362, 374-380, 381-387,
383-389, 387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 157

160/246

FIGURE 158

MDFLLGLCLYWLLRRPSGVVLCLLGACFQMLPAAPSGCPQLCRCEGRLLYCEALNLTEAPHNLSGLLGLSLRY
NSISELRAGQFTGLMQLTWLYLDHNHICSVQGDAFQKLRRVKELTLSSNQITOLPNTTFRPMPNLSVDLSYNK
LQALAPDLFHGLRKLTTLHMRANAIQFVPRIFQDCRSLKFLDIGYNQIKSLARNSFAGLFKLTTELHLEHNDLV
KVNFAHPRLISLHSLCIRRNKVAIVVSSLDWVNLEKMDLSGNEIEYMEPHVFETVPHLQSLQLDNSNRLTYIE
PRILNSWKSLTSITLAGNLWDCGRNVCALASWLSNFQGRYDGNLQCASPEYAQGEDVLDavyAFHLCEDGAEP
SGHLLSAVTNRSDLGPPASSATTIADGGEQHDTFEPATVALPGGEHAENAVQIHKVVTGTMALIIFSFLIVVL
VLYVSWKCFPASLRQLRQCFVTQRRKQKQTMHQMAAMSAQEYYVDYKPNHIEGALVIINEYGSCTCHQQPAR
ECEV

161/246

FIGURE 159

CAGAGAGGAGGCTTGGATTGTCCAGCAGAAACAGAGAAGTCTGAGGTGGTGTCAAGACA
AAAGATGCTTCAGCTTGGAAACTGTTCTCCTGTGCGGCGTGTCACTGGGACCTCAGAGTCT
CTTCTGACAATCTTGGCAATGACCTAACGCAATGTCGTGGATAAGCTGGAACCTGTTCTCA
CGAGGGACTTGAGACAGTTGACAATACTCTTAAAGGCATCCTTGAGAAACTGAAGGTCGACC
TAGGAGTGTTCAGAAATCCAGTGCTTGGCAACTGGCCAAGCAGAAGGCCAGGAAGCTGAG
AAATTGCTGAAACAATGTCATTCTAAGCTGCTTCCAACTAACACGGACATTGGTTGAA
AATCAGCAACTCCCTCATCCTGGATGTCAAAGCTGAACCGATCGATGATGGCAAAGGCCTTA
ACCTGAGCTCCCTGTCACCGCGAATGTCACTGTGGCGGGCCATCATTGGCCAGATTATC
AACCTGAAAGCCTCCTGGACCTCCTGACCGCAGTCACAATTGAAACTGATCCCCAGACACA
CCAGCCTGTTGCCGTCCTGGAGAATGCCAGTGACCCAACCAGCATCTCACTTCCCTGC
TGGACAAACACAGCAAATCATCAACAAGTTCGTGAATAGCGTGATCAACACGCTGAAAGC
ACTGTATCCTCCCTGCTGCAGAAGGAGATATGTCCACTGATCCGCATCTCATCCACTCCCT
GGATGTGAATGTCATTCAAGCAGGTGCGATAATCCTCAGCACAAAACCCAGCTGCAAACCC
TCATCTGAAGAGGAGCAATGAGGAGGACACTGTGGTGCATGCTGATGGTTCCCAGTGGCT
TGCCCCACCCCTTATAGCATCTCCCTCAGGAAGCTGCTGCCACCACTAACCAGCGTGAA
AGCCTGAGTCCCACCGAGAAGGACCTCCCAGATAACCCCTTCCTCACAGTCAGAACAGCAG
CCTCTACACATGTTGTCCTGCCCTGGCAATAAGGCCATTCTGCACCCCTTAA

162/246

FIGURE 160

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59622
><subunit 1 of 1, 249 aa, 1 stop
><MW: 27011, pI: 5.48, NX(S/T): 2
MLQLWKLVLLCGVLTGTSESLLDNLGNDLSNVVDKLEPVLHEGLETVDNTLKGILEKLKV
DLGVLQKSSAWQLAKQKAQEAEKLLNNVISKLPTNTDIFGLKISNSLILDVKAEPIDDG
KGLNLSFPVTANVTVAGPIIGQIINLKASLDLLTAVTIETDPQTHQPVAVLGECASTDPTS
ISLSLLDKHSQIINKFVNSVINTLKVSTVSSLLQKEICPLIRIFIHSLDVNVIQQVVVDNPQ
HKTQLQTLI

Important features:

Signal peptide:

Amino acids 1-15

N-glycosylation sites:

Amino acids 124-128;132-136

N-myristoylation sites:

Amino acids 12-18;16-22;26-32;101-107;122-128;141-147

Leucine zipper pattern:

Amino acids 44-66

163/246

FIGURE 161

CAGCCACAGACGGGTCA**ATGAGCGCGT**ATTACTGCTGGCCCTCTGGGGTTCATCCTCCCAC
TGCAGGAGTGCAGGCCTGCTCTGCCAGTTGGACAGTCAGCATGTGTGGAAGGTGTCC
GACCTACCCGGCAATGGACCCCTAAGAACACCAAGCTGCGACAGCGGCTGGGGTGCAGGA
CACGTTGATGCTCATTGAGAGCGGACCCCAAGTGAAGCCTGGTCTCTCAAGGGCTGCACGG
AGGCCAAGGACCAGGAGCCCCGCTCACTGAGCACCGGATGGGCCCGCTCTCCGTAC
TCCTACACCTTCGTGTGCCAGGAGGACTCTGCAACAAACCTCGTAACTCCCTCCCGCT
TTGGGCCACAGCCCCAGCAGACCCAGGATCCTGAGGTGCCAGTCTGCTTGTCTATGG
AAGGCTGTCTGGAGGGACAACAGAACAGAGATCTGCCCAAGGGACCACACACTGTTATGAT
GCCCTCCTCAGGCTCAGGGAGGAGGCATCTCTCAATCTGAGAGTCCAGGGATGCATGCC
CCAGCCAGGTTGCAACCTGCTCAATGGACACAGGAAATTGGGCCGTGGTATGACTGAGA
ACTGCAATAGGAAAGATTCTGACCTGTCATGGGGGACCACCATTATGACACACGGAAAC
TTGGCTCAAGAACCCACTGATTGGACCACATCGAACATACGAGATGTGCGAGGTGGGGCAGGT
GTGTCAGGAGACGCTGCTCATAGATGTAGGACTCACATCAACCCCTGGTGGGGACAAAAG
GCTGCAGCACTGTTGGGCTCAAAATTCCAGAACGACCACCATCCACTCAGCCCCCTGG
GTGCTTGTGCCCTCTATACCCACTTCTGCTCCTCGAACCTGTGCAATAGTGCCAGCAG
CAGCGTTCTGCTGAACCTCCCTCCTCAAGCTGCCCTGTCCCAGGAGACGGCAGTGT
CTACCTGTGTCAGCCCCCTGGAACCTGTTCAAGTGGCTCCCCCGAATGACCTGCCAGG
GGCGCCACTCATTGTTATGATGGGTACATTGATCTCTCAGGAGGTGGCTGTCCACCAAAAT
GAGCATTCAAGGGCTGCGTGGCCAACCTCCAGCTTGTGAACCACACCAGACAAATCG
GGATCTTCTGCGCGTGAGAACGCGTATGTGCAAGCCTCTCAGCATGAGGGAGGT
GGGGCTGAGGGCCTGGAGTCTCAGTTGGGGGTGGCACTGGCCCCAGCGCTGTG
GTGGGGAGTGGTTGCCCTCCTGCT**TAA**CTCTATTACCCCCACGATTCTCACCGCTGCTGA
CCACCCACACTCAACCTCCCTGACCTCATAACCTAATGGCCTGGACACCAGATTCTTC
CCATTCTGTCATGATCATCTCCCCACACACAATCATTGATCTACTCACCTAACAGCA
ACACTGGGGAGAGCCTGGAGCATCCGGACTTGCCTATGGGAGAGGGGACGCTGGAGGAGTG
GCTGCATGTATCTGATAATACAGACCCCTGTCCTTCA

164/246

FIGURE 162

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59847
><subunit 1 of 1, 437 aa, 1 stop
><MW: 46363, pI: 6.22, NX(S/T): 3
MSAVLLLALLGFILPLPGVQALLCQFGTVQHVWKVSDLPRQWTPKNTSCDSGLGCQDTLM
LIESGPQVSLVLSKGCTEAKDQEPRVTEHRMGPGLSLISYTFVCRQEDFCNNLVNSLPLW
APQPPADPGSLRCPVCLSMEGCLEGTTIEICPKGTTCYDGLLRLRGGGIFSNLRVQGCM
PQPGCNLLNGTQEIGPVGMTENCNRKDFLTCHRGTTIMTHGNLAQEPTDWTTSNTEMCEV
GQVCQETLLLIDVGLTSTLVGKGCSTVGAQNSQKTTIHSAPPGVLVASYTHFCSSDLCN
SASSSSVLLNSLPPQAAAPVPGDRQCPTCQVPLGTCSSGSPRMTCPRGATHCYDGYIHLSG
GGLSTKMSIQGCVAQPPSSFLLNHTRQIGIFSAREKRDVQPPASQHEGGGAEGLESLTWGV
GLALAPALWWGVVCPSC
```

Important features of the protein:**Signal peptide:**

Amino acids 1-15

Transmembrane domain:

Amino acids 243-260

N-glycosylation sites:

Amino acids 46-50;189-193;382-386

Glycosaminoglycan attachment sites:

Amino acids 51-55;359-363

N-myristoylation sites:

Amino acids 54-60;75-81;141-147;154-160;168-174;169-175;
198-204;254-260;261-267;269-275;284-290;333-339
347-353;360-366;361-367;388-394;408-414;419-425

165/246

FIGURE 163

GAGGATTGCCACAGCAGCGGATAGAGCAGGAGAGCACCACCGGAGCCCTTGAGACATCCTTGAGAAGAGCCAC
AGCATAAGAGACTGCCCTGCTTGGTGTGGAGGATGATGGTGGCCCTTCGAGGAGCTCTGCATTGCTGGTT
CTGTTCCCTGAGCTTTCTGCCCGCCGAGCTGACCCAGGACCCAGCATGGTCATTACATCTACCAGCG
CTTTCGAGCTTGAGCAAGGGCTGGAAAAATGTACCCAAGCAACGAGGGCATACATTCAAGAATTCCAAGAGT
TCTAAAAAAATATATCTGTCATGCTGGGAAGATGTCAGACCTACACAAGTGAGTACAAGAGTGCACTGGGTAAC
TTGCCACTGAGAGTTGAACGTGCCAACGGGAGATTGACTACATACAATACCTTGAGAGGCTGACGAGTGAT
CGTATCAGAGGACAAGACACTGGCAGAAATGTTGCTCCAAGAAGCTGAAGAAGAGAAAAAGATCGGACTCTGC
TGAATGCAAGCTGTGACAAACATGCTGATGGGCATAAAAGTCTTGGAAAAATAGTGAAGAAGATGATGGACACAT
GGCTCTTGGATGAAAGATGCTGTCTATAACTCTCCAAAGGTGACTTATTAAATTGGATCCAGAAACAACACTGT
TTGGGAATTGCAAAACATACGGGCATTGAGGAGATAACACCAAGCCAGCTCCCGGAAGCAAAATCTAACAC
TTTCTGGCAGGGAACAGGCAAGTGATCTACAAAGGTTCTATTTCATAACCAAGCAACTTCTAATGAG
ATAATCAAATATAACCTGAGAAGAGGACTGTGAGATGCTGCTCCAGGAGGGTAGGCCAGCATT
GGTTTACCCAGCACTCCCTCAACTTACATTGACCTGGCTGTGGATGAGCATGGGCATCTGGCCATCCACTCTG
GGCCAGGGCACCCATAGCCATTGGTTCTCAGGAAAGATTGAGCCGGGCAACTGGGAGTGAGCATTGATGGGAT
ACCCCATGAGAAGCAGGATGCTGAAGCCTCATCCTCTTGTTCTCTATGTGGTCTACAGTACTGG
GGGCCAGGGCCCTCATGCATCACCTGCATCTATGATCCACTGGGACTATCACTGAGGAGGACTTGCCCAACT
TGTCTTCCCCAAGAGACCAAGAAGACTCCATGATCCATTACAACCCAGAGATAAGCAGCTCTATGCCCTGG
AATGAAGGAAACCAGATCATTACAAACTCCAGACAAAGAGAAAGCTGCTCTGAAGTAATGCAATTACAGCTGT
GAGAAAGAGCACTGTGGCTTGGCAGCTGTTCTACAGGACAGTGAGGCTATAGCCCTTCACAATATAGTATCC
CTCTAATCACACACAGGAAGAGTGTAGAAGTGGAAATACGTATGCCCTCTTCCAAATGTCACTGCCCTAG
GTATCTTCAAGAGCTTAGATGAGAGCATATCATCAGGAAAGTTCAACATGTCATTACTCCCCAAACCTC
CTGGCTCTCAAGGATGACACATTCTGATACAGCCTACTTCAGCCTTTGTTTACTGCTCCCCAGCATTAC
TGTAACTCTGCCATCTCCCTCCCACAATTAGAGTTGTATGCCAGCCCTAATATTCCACACTGGCTTTCTCT
CCCTGGCTTTGCTGAAGCTCTCCCTCTTTCAAAATGTCATTGATATTCTCCATTTCACTGCCCAACT
AAAATACTATTAAATTCTTCTTCTTCTTGGAGAGACAAGGTCTCACTATGTTGCCAGGCTGG
CTCAAACCTCCAGAGCTCAAGAGATCCTCCTGCCCTCAGCCTCTAAAGTACCTGGGATTACAGGCATGTGCCACCA
CACCTGGCTAAAATACTATTCTTATTGAGGTTAACCTCTATTCCCTAGCCCTGCTTCACTAACGCTT
GGTAGATGTAATAATAAGTAAAAATTAAACATTGAAATATCGCTTCCAGGTGTGGAGTGTGTCACATCAT
TGAATTCTCGTTCACCTTGTGAAACATGCACAAGTCTTACAGCTGTCATTCTAGAGTTAGTGAGTAACA
CAATTACAAAGTGAAGAGATACAGCTAGAAAAACTACAAATCCCATAGTTTCCATTGCCCAAGGAAGCATCA
AATACGTATGTTGTTCACCTACTCTTATAGTCATGCTTCAAGCCTTCAATAGGCCCTTCAAAATGATAATTCTCCAGAAAACAGTC
TTAGCCAGTTTCTATGCTGCAAGACCTTCAATAGGCCCTTCAAAATGATAATTCTCCAGAAAACAGTC
TAAGGGTGAGGAGCCCAACTCTAGCCTCTTGTCTGCTCTGCTTTCTGCTTAAATTCA
ATAAAAGTGACACTGAGCAAAAAAAAAAAAAAA

166/246

FIGURE 164

MMVALRGASALLVLFLAAFLLPPPQCTQDPAMVHYIYQRFRVLEQGLEKCTQATRAYIQEFQEF SKNISVMLGRC
QTYTSEYKSAVGNLALRVERAOREIDYIQYLREADECIVSEDKTLAEMLLQEAEKKIRTLLNASCDNMLMGI
KSLKIVKKMMDTHGSWMKDAVYNSPKVYLLIGSRNNTVWEFANIRAFMEDNTKPAPRKQILTLSWQGTGQVIYK
GFLFFHNQATSNEIIKYNLQKRTVEDRMLLPGGVGRALVYQHSPSTYIDLAVDEHGLWAIHSGPGTHSHLVLT
IEPGTGLGVEHSWDTPCRSQDAEASFLLCGVLYVVYSTGGQQGPHRITCIYDPLGTISEEDLPNLFFPKRPRSHSM
IHYNPRDKQLYAWNEGNQIIYKLQTKRKLPLK

167/246

FIGURE 165

TGGCCTCCCCAGCTGCCAGGCACAAGGCTGAGCGGGAGGAAGCGAGAGGCATCTAAGCAGGCAGTGTGTTGCC
TTCACCCCAAGTGA**CCAT**GAGAGGTGCCACCGAGTCTCAATCATGCTCCTCTAGTAACTGTGCTGACTGTG
CTGTGATCACAGGGCCTGTGAGCGGGATGTCCAGTGTGGGCAGGCACCTGCTGTGCCATCAGCCTGTGGCTT
CGAGGGCTGCGGATGTGACCCCGCTGGGCGGGAAAGCGAGGAGTGCCACCCCGCAGCCACAAGGTCCCCCTT
CTTCAGGAAACGCAAGCACCACACCTGTCCTGCTGCCAACCTGCTGTGCTCCAGGTTCCCGGACGGCAGGT
ACCGCTGCTCATGGACTGAAAGAACATCAATTTC**AGGCGCTTGCC**TGCTCAGGATAACCCACCATCCTTT
CTGAGCACAGCCTGGATTTTATTCTGCCATGAAACCCAGCTCCCAGTACTCTCCAGTCCCTACACTGACT
ACCTGTATCTCTTGCTAGTACGCACATATGCACACAGGCAGACATACTCCCATCATGACATGGTCCCCAG
GCTGGCCTGAGGATGTACAGCTTGAGGCTGTGGTGTGAAAGGTGGCAGCCTGGTCTCTCCCTGCTCAGGC
TGCCAGAGAGGTGGTAAATGGCAGAAAGGACATCCCCCTCCCTCCCCAGGTGACCTGCTCTCTTCCGGC
CCTGCCCCCTCCTCCCCACATGTATCCCTCGGTCTGAATTAGACATTCCCTGGGCACAGGCTCTGGGTGCATTGCT
CAGAGTCCCAGGTCCCTGGCTGACCTCAGGCCCTCACGTGAGGTCTGTGAGGACCAATTGTGGGTAGTTCA
TCTCCCTCGATTGGTTAACTCCTTAGTTCAAGACACAGACTCAAGATTGGCTCTCCAGAGGGCAGCAGAC
AGTCACCCCAAGGCAGGTGTAGGGAGGCCAGGGAGGCAATCAGCCCCCTGAAGACTCTGGTCCAGTCAGCCT
GTGGCTTGTGGCTGTGACCTGTGACCTCTGCCAGAATTGTCATGCCCTGTGAGGCCCTCTTACACACTTT
ACCACTTAACCAACTGAAGCCCCCAATTCCACAGCTTCCATTAAAATGCAAATGGTGGTCAATCTAAT
CTGATATTGACATATTAGAAGGCAATTAGGGTGTTCCTAAACAACCTCTTCAAGGATCAGCCCTGAGAGC
AGGGTGGTGA~~CTT~~GAGGAGGGCAGTCCTGTCCAGATTGGGTGGGAGCAAGGGACAGGGAGCAGGGCAGGG
GCTGAAAGGGCAGTCAGACCAAGGGAGGCAACTACACACCAACATGCTGGCTT~~AGA~~ATAAAAGCACCAA
CTGAAAAAA

168/246

FIGURE 166

MRGATRVSIMLLLTVSDCAVITGACERDVQCGAGTCCAISLWLRGLRMCTPLGREGECHPGSHKVPFFRKHK
HHTCPCLPNLLCSRFPDGRYRCSMDILKNINF

Important features:

Signal peptide:

amino acids 1-19

Tyrosine kinase phosphorylation site:

amino acids 88-95

N-myristoylation sites:

amino acids 33-39, 35-41, 46-52

169/246

FIGURE 167

AACTCAAACCTCTCTGGGAAAACGCGGTGTTGCTCTCCCGGAGTGGCCTTGGCAGGGTGTGGAGCCC
TCGGTCTGCCCGTCCGGTCTCTGGGCCAAGGCTGGTTCCCTCATGTATGGCAAGAGCTACTCGTGC
TGCCTCTCTCCCTGGCATAACAGCTCACAGCTCTGGCTATAGCAGCTGTGAAATTATACCTCCC
CTGGAGGCTGTTAATGGGACAGATGCTCGGTTAAATGCACCTTCTCCAGCTTGCCCTGTGGGTGATG
AACAGTGACCTGGAATTTCGTCTAGACGGGGACCTGAGCAGTTGTATTCTACTACCACATAGAT
TCCAAACCCATGAGTGGCGGTTAAGGACCGGGTGTCTGGATGGGAATCCTGAGCGGTACGATGC
CTTCTCTGGAAACTGCAGTCGACGACAATGGGACATAACCTGCCAGGTGAAGAACCCACCTG
GGTGTAGGGAGATCCGGCTCAGCGTGTGACACTGTACGCTCTGTAGATCCACTTCTGGCTCTGG
TTGGCTCTGCCGTGTGACTGATGATCATATAAGTAATTGTAGTGGCCTTCCAGCATTACCG
TGGCGAAAGAGCTCATAAAGTGGGGAGATAAAATCAAAGAAGAGGCTCAACCAAGAGAAAAGGT
CTCTGTTATTAGAACACAGACTAACAAATTAGATGGAAGGCTGAGATGATTTCAAGAACAA
GTATTTCTTGAAGTTAATGGAAACTTTCTTGCTTTCCAGTTGTGACCCGTTTCCAACCAG
ATATTAGATTCTAGACAAGCAACACCCCTCTGGAGCCAGCACAGTGTCTCCATATC
CTCATTATTAAAGGTCTTATTAAATTCAAGGTGAAATTTCAGAGTGTAAATT
CTACATTGGCTTAAGACACTACTACAGTGTATGACTGTATA
AAAGCCAATTGTCTGTACATTCCCTTCACGTATTCTTAGCAGCA
TTACTCTCTTCCCTCCACATTCTCAATTAAAGGTGAGCTAAGC
TCCTAAATTCAAACGTAAATGACATT
TGAATTCTTCAATTCCAGGTGATAGATTTCG

170/246

FIGURE 168

MYGKSSTRAVLLLLGIQLTALWPIAAVEIYTSRVLEAVNGTDARLKCTFSSFAPVGDALTVTWNFRPLDGGEQ
FVFFYHIDPFQPMMSGRFKDRVSDGNGPERYDASILLWKLQFDDNGTYTCQVKNPPDGVIGEIRLSVVHTVRF
SEIHFLALAIGSACALMIIIVVVVLFQHYRKWRWAERAHKVVEIKSKEEERLNQEKKVSVYLEDTD

171/246

FIGURE 169

GAGCGAACATGGCAGCGCGTGGCGGTTGGTGTCTCTGTGACCATGGTGGTGGCGCTG
CTCATCGTTGCGACGTTCCCTCAGCCTCTGCCAAAGAAAAGAAGGAGATGGTGTATCTGA
AAAGGTTAGTCAGCTGATGGAATGGACTAACAAAGACCTGTAATAAGAATGAATGGAGACA
AGTTCGTCGCTTGTGAAAGCCCCACCGAGAAATTACTCCGTTATCGTCATGTTCACTGCT
CTCCAAC TGCA TAGACAGTGTGCGTTGCAAGCAAGCTGATGAAGAATTCCAGATCCTGGC
AAACTCCTGGCGATACTCCAGTCATTACCAACAGGATATTGGCCATGGTGGATTG
ATGAAGGCTCTGATGTATTCAGATGCTAACATGAATTAGCTCCAACTTCATCAACTTT
CCTGCAAAAGGGAAACCCAAACGGGGTGATACATATGAGTTACAGGGTGCAGGGGTTTCAGC
TGAGCAGATTGCCCGGTGGATGCCGACAGAACTGATGTCATATTAGAGTGATTAGACCC
CAAATTATGCTGGTCCCCTTATGTTGGGATTGCTTTGGCTGTTATTGGTGGACTTGTAT
CTTCGAAGAAGTAATATGGAATTCTCTTTAATAAAACTGGATGGCTTGCAGCTTGTG
TTTGTGCTTGTCTGATGACATCTGGTCAAATGTGAAACCATAAAGAGGACCACCATATGCC
ATAAGAATCCCCACACGGGACATGTGAATTATATCCATGGAAGCAGTCAGCCCAGTTGTA
GCTGAAACACACATTGTTCTGTGTTAATGGTGGAGTTACCTTAGGAATGGTGTCTTATG
TGAAGCTGCTACCTCTGACATGGATATTGGAAAGCGAAAGATAATGTTGTGGCTGGTATTG
GACTTGTGTTGATTATCTTCAGTTGGATGCTCTATTAGATCTAAATATCATGGCTAC
CCATACAGCTTCTGATGAGTAAAAAAGGTCCCAGAGATATAGACACTGGAGTACTGGAA
ATTGAAAAACGAAAATCGTGTGTTGAAAAGAAGAATGCAACTTGTATATTGTATTAC
CTCTTTTTCAAGTGTATTAAATAGTTAACCAAAGAAGATGTGTAGTGCCTTA
ACAAGCAATCCTCTGTCAAAATCTGAGGTATTGAAAATAATTATCCTCTTAACCTCT
CCCACTGAACTTTATGGAACATTAAATTAGTACAATTAAAGTATATTATAAAATTGTAAAA
CTACTACTTGTGTTAGTTAGAACAAAGCTAAAACACTTTAGTTAACCTGGTCTGAT
TTTATATTGCCTTATCCAAAGATGGGAAAGTAAGTCTGACAGGTGTTCCACATATGCC
TGTACAGATAACTACATTAGAACATTCTAGCTTCTCATCTTGTGTGGATGTGTAT
ACTTACGCATCTTCCTTTGAGTAGAGAAATTATGTGTGTCTGTTCTGAAATG
GAACACCATTCTCAGAGCACACGTCTAGCCCTCAGCAAGACAGTTGTTCTCCTCCT
GCATATTCTACTGCGCTCCAGCCTGAGTGATAGAGTGAGACTCTGTCCTAAAAAGTA
TCTCTAAATACAGGATTATAATTCTGCTTGAGTAGGTGTTAACTACCTGTATTAGAAA
GATTTCAGATTCACTCCATCTCCTTAGTTCTTAAAGGTGACCCATCTGTGATAAAAATA
TAGCTTAGTGCTAAATCAGTGTAACTTATACATGGCTAAAATGTTCTACAAATTAGAGT
TTGTCACTTATTCCATTGTACCTAACAGAGAAAAATAGGCTCAGTTAGAAAAGGACTCCCTGG
CCAGGCGCAGTGACTTACGCCCTGTAATCTCAGCACTTGGGAGGCCAAGGCAGGCAGATCAC
GAGGTCAAGGAGTTGAGACCACCTGGCCAACATGGTGAACACCCGCTCTACTAAAATAT
AAAAATTAGCTGGGTGTGGTGGCAGGAGCCTGTAATCCAGCTACACAGGAGGCTGAGGCAC
GAGAATCACTGAACTCAGGAGATGGAGGTTTCAGTGAGCCGAGATCACGCCACTGCACTCC
AGCCTGGCAACAGAGCGAGACTCCATCTAAAAAA

172/246

FIGURE 170

MAARWRFWCVSVMVALLIVCDVPSASAQRKKEMVLSEKVSQLMEMTNKRPVIRMNGDKFR
RLVKAPPRNYSVIVMFTALQLHRQCVVCKQADEEFQILANSWRYSSAFTNRIFFAMVDFDEG
SDVFQMLNMNSAPTFINFPAKGPKRGDTYELQVRGFSAEQIARWIADRTDVNIRVIRPPNY
AGPLMLGLLLAVIGGLVYLRRSNMEFLFNKTGWAFAALCFVLAATSGQMWNHIGPPYAHKN
PHTGHVNYIHGSQAFVAETHIVLLFNGGVTLMVLLCEAATSDMDIGKRKIMCVAGIGLV
VLFFSWMLSIFRSKYHGYPYSLMS

Signal peptide:
amino acids 1-29

Transmembrane domains:
amino acids 183-205, 217-237, 217-287, 301-321

173/246

FIGURE 171

CTCCACTGCAACCACCCAGAGCCATGGCTCCCCGAGGCTGCATCGTAGCTGTCTTGCCATTTCTGCATCTCC
AGGCTCCTCTGCTCACACGGAGCCCCAGTGGCCCCCATGACTCCTTACCTGATGCTGTGCCAGCACACAAGAG
ATGTGGGGACAAGTTCTACGACCCCCCTGCAGCACTGTTGCTATGATGATGCCGTCGTGCCCTTGGCCAGGACCC
AGACGTGTGGAAACTGCACCTTCAGAGTCTGCTTGAGCAGTGCTGCCCTGGACCTTCATGGTGAAGCTGATA
AACCGAGAACTGCGACTCAGCCCCGGACCTCGGATGACAGGCTTGTCGAGTCAGCTAATGGAACATCAGGGG
AACGATGACTCCTGGATTCTCCTTCCCTGGGCTGGAGAAAGAGGCTGGTGTACCTGAGATCTGGGATGC
TGAGTGGCTGTTGGGGCCAGAGAAACACACACTCAACTGCCCACTTCATTCTGTGACCTGTCTGAGGCCAC
CCTGCAGCTGCCCTGAGGAGGCCACAGGTCCCCCTCTAGAATTCTGGACAGCATGAGATGCGTGTGCTGATGG
GGGCCAGGGACTCTGAACCCCTCTGATGACCCCTATGGCCAACATCAACCCGGCACCACCCCAAGGCTGGCTG
GGGAACCCCTCACCCCTCTGAGATTTCATCTCAAGTTCTTCTATCCAGGAGCAAAGCACAGGATC
ATAATAAATTATGTAATTATAAATGAAAA

174/246

FIGURE 172

MAPRGCIVAVFAIFCISRLLC SHGAPVAPMTPYLMLCQPHKRCGDKFYDPLQHCCYDDAVVPLARTQTCGNCTF
RVCFEQCCPWTFMVKLINQNCDSARTSDDRLCRSVS

Important features:

Signal peptide:

amino acids 1-24

175/246

FIGURE 173

GGGGCGGGTGCCTGGAGCACGGCGCTGGGGCCGCCGCAGCGCTCACTCGCTCGCACTCAG
TCGCAGGGAGGCTTCCCCCGCGCCGGCGTCCCAGCGCTCCCCGGCACCAGAAGTCCCTCT
GCGCGTCCGACGGCGACATGGCGTCCCCACGGCCCTGGAGGCCAGCTGGCGCTGGGA
TCCCTGCTCTCGCTCTTCCCTGGCTGCCTAGGTCCGGTGGCAGCCTCAAGGTGC
CACGCCGTATTCCCTGTATGTCGTCCCAGGGCAGAACGTCACCCACCTGCAGGCTCT
TGGGCCCTGTGGACAAAGGGCAGATGTGACCTCTACAAGACGTGGTACCGCAGCTCGAGG
GGCAGGGTGCAGACCTGCTCAGAGCGCCGGCCATCCGCAACCTCACGTTCCAGGACCTTCA
CCTGCACCATGGAGGCCACCAGGCTGCCAACACCAGCCACGACCTGGCTCAGGCCACGGC
TGGAGTCGGCCTCCGACCACCATGGCAACTTCTCCATCACCATGCGCAACCTGACCCCTGCTG
GATAGCGGCCTCTACTGCTGCCTGGTGGAGATCAGGCACCACACTCGGAGCACAGGGT
CCATGGTGCCATGGAGCTGCAGGTGCAGACAGGGCAAAGATGCACCATCCAACACTGTGTGGTGT
ACCCATCCTCCTCCCAGGATAGTGAACATCACGGCTGCAGCCCTGGCTACGGGTGCCTGC
ATCGTAGGAATCCTCTGCCTCCCCCTCATCCTGCTCCTGGTCTACAAGCAAAGGCAGGCAGC
CTCCAACCGCCGTGCCAGGAGCTGGTGCAGACAGCAACATTCAAGGGATTGAAAACC
CCGGCTTGAGCCTCACCACTGCCAGGGATAACCGAGGCCAAAGTCAGGCACCCCTG
TCCTATGTGGCCAGCGGCAGCCTCTGAGTCTGGCGGCATCTGCTTCGGAGCCCAGCAC
CCCCCTGTCTCCTCCAGGCCCGGAGACGTCTTCTCCATCCCTGGACCCCTGCTGACT
CTCCAAACTTGAGGTACAT**TAG**CCCAGCTGGGGACAGTGGGCTGTGTGGCTGGGTCTGG
GGCAGGTGCATTGAGCCAGGGCTGGCTCTGTGAGTGGCCTCTGGCTCGGCCCTGGTTC
CCTCCCTCTGCTCTGGCTCAGATACTGTGACATCCCAGAAGCCCAGGCCCTCAACCCCTC
TGGATGCTACATGGGATGCTGGACGGCTCAGCCCTGTTCCAAGGATTTGGGTGCTGAG
ATTCTCCCTAGAGACCTGAAATTCAACCAGCTACAGATGCCAAATGACTTACATCTTAAGAA
GTCTCAGAACGTCCAGCCCTCAGCAGCTCTCGTCTGAGACATGAGCCCTGGGATGTGCA
GCATCAGTGGGACAAGATGGACACTGGGCCACCCCTCCAGGCACAGACAGGGCACGGTG
GAGAGACTCTCCCCCGTGGCCGCTTGGCTCCCCGTTTGCCGAGGCTGCTCTGTG
AGACTTCCTTTGTACCACAGTGGCTCTGGGCCAGGCCTGCCTGCCACTGGCCATGCC
ACCTCCCCAGCTGCCTCCTACCAGCAGTTCTGAAGATCTGTCAACAGGTTAAGTCAAT
CTGGGGCTTCCACTGCCTGCATTCCAGTCCCCAGAGCTGGTGGTCCCAGGGAAAGTAC
ATATTGGGCATGGTGGCCTCCGTGAGCAAATGGTGTCTGGGCAATCTGAGGCCAGGACAG
ATGTTGCCCAACCACTGGAGATGGTGTGAGGGAGGTGGTGGGCCTCTGGGAAAGGTGA
GTGGAGAGGGGACCTGCCCTCCCCGCCCTCCCTACTCCACTGCTCAGCGCGGGCC
ATTGCAAGGGTGCACACAATGTCTTGTCCACCTGGACACTTCTGAGTATGAAGCAGGGAT
GCTATTAAAAACTACATGGGGAAAAAAAAAAAAAAAAAAAAAAAAGA

176/246

FIGURE 174

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64897
><subunit 1 of 1, 311 aa, 1 stop
><MW: 33908, pI: 6.87, NX(S/T): 6
MGVPTALEAGSWRWGSLLFALFLAASLGPVAAFKVATPYSLYVCPEGQNVTLTCRLLGPVDK
GHDVTFYKTWYRSSRGEVQTCSEERRPIRNLTFQDLHLHHGGHQAANTSHDLAQRHGLEASD
HHGNFSITMRNLTLDSGLYCCCLVVEIRHHHSEHRVHGAMELQVQTGKDAPSNCVVPSSQ
DSENITAAALATGACIVGILCLPLILLLVYKQRQAASNRRAQELVRMDSNIQGIENPGFEAS
PPAQGIPEAKVRHPLSYVAQRQPSESRHLLSEPSTPLSPPPGDVFFPSLDPPDSPNFEVI
```

Signal peptide:
amino acids 1-28

Transmembrane domain:
amino acids 190-216

177/246

FIGURE 175

178/246

FIGURE 176

MDSLRKMLISVAMLGAGAGVGYALLVIVTPGERRKQEMLKEMPLQDPRSREEAARTQQLLLATLQEAAATTQENV
AWRKNWMVGGEGGASGRSP

Important features:

Signal peptide:

amino acids 1-18

179/246

FIGURE 177

GCAGGCAGGTGGCCTCAGGAGGTGCCTCCAGGCAGGCCAGTGGCCTGAGGCCAGCAAG
GGCTAGGGTCCATCTCAGTCCCAGGACACAGCAGCAGGCCACCATGGCCACGCCCTGGCTCC
AGCAGCATCAGAGCAGCCCTGTGGTGTGGCAGCAAAGTTAGCTGGCTGGCCCGCTGTGA
GGGGCTTCGCGCTACGCCCTGCGGTGTCCCAGGGCTGAGGTCTCCTCATCTCCCTAGC
AGTGGATGAGCAACCCAACGGGGCCGGGAGGGAACTGGCCCCGAGGGAGAGGAACCC
AAAGCCACATCTGTAGCCAGGATGAGCAGTGTGAATCCAGGCAGCCCCCAGGACCAGGGAGG
CACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCTGCCCTGTCCGGGGATGACTGATT
TCCTCCGCCAGGCCACCCAGAGGAGAACGGCACCCGCCCTGGAGGCACAGGCCATGAGGGC
TCTCAGGAGGTGCTGCTGATGTGGCTCTGGTGTGGCAGTGGCGGCACAGACACGCC
CCGGCCCGGCCGTAGGGTGTGTGCTGCCGGCTCACGGGACCCCTGCTCCGAGTCGTTG
TGCAGCGTGTGTACCAAGCCCTCCTCACACCTGCGACGGGCACCGGGCCTGCAGCACCTAC
CGAACCATCTATAGGACGCCATCCGCCAGGCCCTGGGCTGGCCCTGCCCCCTGCCAGGCCTCGCTA
CGCGTGTGCCCGGCTGGAAGAGGACCAGCGGGCTTCCTGGGCTGGCCCTGTGGAGCAGCAATAT
GCCAGCCGCCATGCCGAACGGAGGGAGCTGTGTCCAGCCTGGCCGTGCCCTGCC
GGATGGCGGGGTGACACTGCCAGTCAGATGTGGATGAATGCACTGCTAGGAGGGCGGCTG
TCCCCAGCGCTGCATCAACACCGCCGGCAGTTACTGGTGCCAGTGTGGAGGGCACAGCC
TGTCTGCAGACGGTACACTCTGTGTGCCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCG
ACAGGAGTGGACAGTGAATGAAGGAAGAAGTGCAGAGGCTGCAGTCAGGTTGGACCTGCT
GGAGGAGAAGCTGCAGCTGGTCTGCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGC
ATGGGCTCCGGACCCCGCAGCCTGGTGCACTCCTCCAGCAGCTGGCCGCATCGAC
TCCCTGAGCGAGCAGATTCTCCTGGAGGAGCAGCTGGGCTCTGCTCTGCAAGAAAGA
CTCGTGACTGCCAGCGCTCCAGGCTGGACTGAGCCCTCACGCCGCCCTGCAGCCCCCATG
CCCCCTGCCAACATGCTGGGGTCCAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGC
AGGGCCTCCTCCTCTCCTCCCTGGCTACCCCTGGCTACCCCAACCCCTGGCTACCCCAACGGCA
GGGCTGGGATCTCTGTGAATCCACCCCTGGCTACCCCAACCCCTGGCTACCCCAACGGCA
TCCAAGGCCAGGTGGACCCCTCAGCTGAGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGAC
CCATGGCACAGGCCAGGCAGCCGGAGGCTGGGTGGGCCTCAGTGGGGCTGCTGCCTGAC
CCCCAGCACAATAAAATGAAACGTG

180/246

FIGURE 178

MRGSQEVL LMWLLV LAVGGTEHAYRPGR RVCAVRAH GDPVSE SFVQRVYQPFLTTCDGH RAC
STYRTIYRTAYRRSPGLAPARPRYACCPG WKRTS GLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTCQSDVDECSARRGGCPQRCINTAGSYWCQCWE GHLSADGTL CVPKG GPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKDS

Signal sequence:

1-19

181/246

FIGURE 179

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTGCCCTCGCTACGCAG
AGCCTCTCCGTGGCTCCGCACCTTGAGCATTAGGCCAGTCTCCTCTCTCTAATCCAT
CCGTACACCTCTCCTGTCATCCGTTCCATGCCGTGAGGTCCATTACAGAACACATCC**ATGG**
CTCTCATGCTCAGTTGGTTCTGAGTCTCCTCAAGCTGGATCAGGGCAGTGGCAGGTGTT
GGCCAGACAAGCCTGTCCAGGCCCTGGTGGGGAGGACGCAGCATTCTCCTGTTCTGTC
TCCTAAGACCAATGCAGAGGCCATGGAAGTGCGGTCTTCAGGGCCAGTCTAGCGTGG
TCCACCTCTACAGGGACGGGAAGGACCAGCATTATGCAGATGCCACAGTATCAAGGCAGG
ACAAAACGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTGAGGCTGGAAAACATTAC
TGTGTTGGATGCTGCCCTATGGGTGAGGATTAGTCCCAGTCTTACTACCAGAACGGCA
TCTGGGAGCTACAGGTGTCAGCACTGGGCTCAGTCTCCTCATTTCCATCACGGGATATGTT
GATAGAGACATCCAGCTACTCTGTCAGTCCTCGGGCTGGTCCCCCGGCCACAGCGAAGTG
GAAAGGTCCACAAGGACAGGATTGTCACAGACTCCAGGACAAACAGAGACATGCATGGCC
TGTGTTGATGTGGAGATCTCTGACCGTCCAAGAGAACGCCGGAGCATACTCTGTTCCATG
CGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAGGAGATACCTTTTCA
GCCTATATCGTGGCACCTGGTACCAAAGTACTGGAATACTCTGCTGTGGCTATTTTG
GCATTGTTGGACTGAAGATTCTTCTCCAATTCCAGTGGAAAATCAGGCGGAACGGAC
TGGAGAAGAACAGCAGGACAGGCAGAATTGAGAGACGCCGGAAACACGCAGTGGAGGTGAC
TCTGGATCCAGAGACGGCTACCCGAAGCTCTGCGTTCTGATCTGAAAACGTAAACCCATA
GAAAAGCTCCCAGGAGGTGCCTCACTCTGAGAAGAGATTACAAGGAAGAGTGTGGTGGCT
TCTCAGAGTTCCAAGCAGGGAAACATTACTGGGAGGTGGACGGAGGACACAATAAGGTG
GCCGTGGAGTGTGCCGGATGATGTGGACAGGAGGAAGGAGTACGTGACTTGTCTCCG
ATCATGGGTACTGGGTCTCAGACTGAATGGAGAACATTGTATTCACATTAATCCCCT
TTATCAGCGTCTTCCCAGGACCCCACCTACAAAAATAGGGTCTTCCCTGGACTATGAGTG
TGGGACCATCTCCTCTTCAACATAATGACCAGTCCCTATTATACCCCTGACATGTCGGT
TTGAAGGCTTATTGAGGCCCTACATTGAGTATCCGCTCTATAATGAGCAAAATGGAACCTCC
ATAGTCATCTGCCAGTCACCCAGGAATCAGAGAAAGAGGCCCTTGGCAAAGGGCTCTGC
AATCCCAGAGACAAGAACAGTGAGTCCTCCTCACAGGCAACCACGCCCTCCCTCCAGGG
GTGAAATGTAGGATGAATCACATCCCACATTCTTCTTAGGGATATTAAGGTCTCTCCCA
GATCCAAAGTCCCGCAGCAGCCGGCAAGGTGGCTCCAGATGAAGGGGACTGGCCTGTCC
ACATGGGAGTCAGGTGTCTGGCTGCCCTGAGCTGGAGGAAGAAGGCTGACATTACATT
AGTTGCTCTCACTCCATCTGCTAAGTGATCTGAAATACCACCTCAGGTGAAGAACCG
TCAGGAATTCCCACAGGCTGTGGTAGATTAAGTAGACAAGGAATGTGAATAATGC
TTAGATCTTATTGATGACAGAGTGATCCTAATGGTTGTCATTATATTACACTTCAGTA
AAAAAA

182/246

FIGURE 180

MALMLSIVSLLKLGGQWQVFGPDKPVQALVGEDAASFCLSPKTNAEAMEVRFFRGQFSS
VVHLYRDGKDQPFMOMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISQSYYQK
AIWELOVSALGSVPLISITGYVDRDIQLCQSSGWPRPTAKWKGPGQGDLSTDRTNRDMH
GLFDVEISLTQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLF
FGIVGLKIFFSKFQWKIQAELDWRRKHGQAEELRDARKHAVEVTLDPETAHPKLCVSDLKTVT
HRKAPQEVPHSEKRFRKSVVASQSFQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLS
PDHGYWVRLNGEHLYFTLNPRFISVFPPRKIGVFLDYECGTISFFNINDQSLIYTLTC
RFEGLLRPYIEYPSYNEQNGTPIVICPVTQESEKEASWQRASAIPETSNSSESSSQATPF
RGEM

Signal peptide:
amino acids 1-17

Transmembrane domain:
amino acids 239-255

183 / 246

FIGURE 181

184/246

FIGURE 182

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64952
><subunit 1 of 1, 258 aa, 1 stop
><MW: 25716, pI: 8.13, NX(S/T): 5
MRSLPSLGGLALLCCAAAAAAVASAASAGNVTGGGAAAGQVDASPGPGLRGEPSHPFPRATA
PTAQAPRTGPPRATVHRPLAATSPAQSPTTPLWATAGPSSTTFQAPLGPSPTPAAERTS
TTSQAPTRPAPTTLS TTGPAPTPVATTVPAPTPRTPTPDLPSNSSVLPTPPATEAPS
SPPPEYVCNCVVGSLNVNRCNQTTGQCECRPGYQGLHCETCKEGFYLNYSGLCQPCDCSP
HGALSIPCNR
```

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 30-33, 172-175, 195-198, 208-211, 235-238

EGF-like domain cysteine pattern signature.

amino acids 214-226.

185/246

FIGURE 183

TGCGGCGCAGTGTAGACCTGGGAGGATGGCGGCCTGCTGCTGGCTGCTTTCTGGCTTGTCGGTGCCA
GGGCCAGGCCGTGTGGTGGGAAGACTGGACCCCTGAGCAGCCTTGTGGCCCTGGTACGTGCTTGCCTGGCC
TCCCGGGAAAAGGGCTTGCATGGAGAAGGACATGAAGAACGTCGTGGGGGTGGTGGTACCCCTCACTCCAGA
AAACAACCTGCGGACGCTGCTCTCAGCACGGCTGGGAGGGTGTGACCAGAGTGTACGGACTGATAAAGC
GAAACTCCGGATGGGTGTTGAGAATCCCTCAATAGGCCTGCTGGAGCTGGGTGCTGCCACCAACTTCAGA
GACTATGCCATCATCTTCACTCAGCTGGAGTTCGGGGACGAGCCCTAACACCGTGGAGCTGTACAGTCTGAC
GGAGACAGCCAGCCAGGAGGCCATGGGCTCTCACCAAGTGGAGCAGGAGCCTGGGCTTCGTACAGTAGC
AGGCCAGCTGCAGAAGGACCTCACCTGTGCTACAAGATCTTCTGTGAGTGCTGCGTCCCCAGTAGGGATGG
CGCCCACAGGGCTGTGACCTCGGCCAGTGTCCACCCACCTCGCTCAGCGGCTCCCAGCACCAGCT
CAGAATAAAGCGATTCCACAGCA

186/246

FIGURE 184

MGGLLLAALVALSVPRAQAVWLGRILDPEQLLGPWYVLAVASREKGFAMEKDMKNVVGVVVTLTPENNLRTLSS
QHGLGGCDQSVMDLIKRNSGWVFENPSIGVLELWVLATNFRDYAIIFTQLEFGDEPFNTVELYSLTETASQEAM
GLFTKWSRSLGFLSQ

Important features:

Signal peptide:

amino acids 1-20

187/246

FIGURE 185

GTTCGCAGATGCAGAGGTTGAGGTGGCTGCGGACTGGAAGTCATCGGGCAGAGGCTCTCACAGCAGCCAAGGA
ACCTGGGGCCCGCTCCCTCCCCCTCCAGGCCATGAGGATTCTGAGTTAATCTGCTCTGGCAACAGGGC
TTGTAGGGGGAGAGACCAGGATCATCAAGGGGTTCGAGTGCAAGCCTCACTCCCAGCCCTGGCAGGCAGCCCTG
TTCGAGAAGACGCCGGCTACTCTGTGGGGCAGCCTACGCCCTCAGATGGCTCTGACAGCAGCCCAGTGCCT
CAAGCCCCGCTACATAGTTACCTGGGGCAGCACAAACCTCCAGAAGGAGGGCTGTGAGCAGACCCGGACAG
CCACTGAGTCCTTCCCCCACCCGGCTTCAACAAACGCCCTCCCCAACAAAGACCACCGCAATGACATCATGCTG
GTGAAGATGGCATGCCAGTCTCCATCACCTGGGTGTGCGACCCCTCACCTCAGCTGTGTCAGTGC
TGGCACCAGCTGCTCATTTCCGGCTGGGCAGCACGTCCAGCCCCAGTTACGCCCTGCTCACACCTTGCAG
GCGCAACATCACCATCATTGAGCACCAGAAGTGTGAGAACCCCTACCCGGCAACATCACAGACACCATGGTG
TGTGCCAGCGTGCAGGAAGGGGCAAGGACTCCTGCCAGGGTGACTCCGGGGCCCTCTGGTCTGTAACCAGTC
TCTTCAAGGCATTATCTCTGGGCCAGGATCCGTGTGCGATCACCGAAAGCCTGGTGTCTACAGAAAGTCT
GCAAATATGTGGACTGGATCCAGGAGACGATGAAGAACAAATTAGACTGGACCCACCCACCAGCCCATCACCC
TCCATTTCACCTGGTGTGGTTCTGTTACTCTGTTAAATAAGAACCTAACGCAAGACCCCTACGAACA
TTCTTTGGGCCTCCTGGACTACAGGAGATGCTGTCACCTAAATAATCAACCTGGGTTGAAATCAGTGAGACT
GGATTCAAATTCTGCCTGAAATATTGTGACTCTGGGAATGACAACACCTGGTTGTTCTGTTGATCCCCA
GCCCAAAGACAGCTCCTGGCCATATATCAAGGTTCAATAAAATATTGCTAAATGAAAAAAA
AAAAAAAAAAAAAAAAAAAAA

188/246

FIGURE 186

MRILQLILLALATGLVGETRIIKGFECKPHSQWQAALEKTRLLCGATLIAPRWLLTAHCLKPRYIVHLGQ
HNQKEEGCEQTRTATESFPHPGFNNSLPNKDHRNDIMLVKMASPVSITWAVRPLTLSSRCVTAGTSCLISGWG
STSSPQLRLPHTLRCANITIIIEHQKCENAYPGNITDTMVCASVQEGGKDSCQGDGGPLVCNQSLQGIISWGQD
PCAITRKPGVYTKVCKYVDWIQETMKNN

Important features:

Signal peptide:

amino acids 1-18

Serine proteases, trypsin family, histidine active site.
amino acids 58-63

N-glycosylation sites.

amino acids 99-102, 165-168, 181-184, 210-213

Glycosaminoglycan attachment site.

amino acids 145-148

Kringle domain proteins.

amino acids 197-209, 47-64

Serine proteases, trypsin family, histidine protein
amino acids 199-209, 47-63, 220-243

Apple domain proteins

amino acids 222-249, 189-222

189/246

FIGURE 187

GCTCAAGTGCCTGGCCCTGCCTTGCCTGGCCCCACCCAGCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGC
TCTCTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGCGCTGGAGGGCTGTCCTGC
ACCATGGTCCCTGCCTGGCTGTGGCTGCTTGTCTCCGTCCCCCAGGCTCTCCCCAAGGC
CCAGCCTGCAGAGCTGTCTGTGGAAGTCCAGAAAATATGGTGGAAATTCCCTTATACC
TGACCAAGTTGCCGTGCCCCGTGAGGGGGCTGAAGGCCAGATCGTCTGTCAGGGACTCA
GGCAAGGCAACTGAGGGCCCATTGCTATGGATCCAGATTCTGGCTTCCTGCTGGTGACCAG
GGCCCTGGACCGAGAGGAGCAGGCAGAGTACCGCTACAGGTACCCCTGGAGATGCAGGATG
GACATGTCTTGTGGGGTCCACAGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTG
CCCCATTCTCTCAAGCCATCTACAGAGCTCGGCTGAGCCGGGGTACCAAGGCCTGGCATCCC
CTTCCTCTTCCCTGAGGCTTCAGACCGGGATGAGCCAGGCACAGCCAACTCGGATCTCGAT
TCCACATCCTGAGCCAGGCTCCAGCCCCAGCCTTCCCCAGACATGTTCCAGCTGGAGGCTCGG
CTGGGGCTCTGGCCCTCAGCCCCAAGGGGAGCACAGCCTTGACCACGCCCTGGAGAGGAC
CTACCAGCTTGGTACAGGTCAAGGACATGGGTGACCAGGCCTCAGGCCACCCAGGCCACTG
CCACCGTGGAAAGTCTCATAGAGAGCACCTGGGTGTCCTAGAGCCTATCCACCTGGCA
GAGAATCTCAAAGTCCATACCCGCACCACATGGCCCAGGTACACTGGAGTGGGGGTGATGT
GCACTATCACCTGGAGAGCCATCCCCCGGGACCCCTTGAAGTGAATGCAAGGGAAACCTCT
ACGTGACCAGAGAGCTGGACAGAGAAGGCCAGGCTGAGTACCTGCTCCAGGTGCGGGCTCAG
AATTCCCATGGCGAGGACTATGCGGCCCTCTGGAGCTGCACGTGCTGGTGTGAGAA
TGACAAACGTGCCTATCTGCCCTCCCCGTGACCCCACAGTCAGCATCCCTGAGCTAGTCCAC
CAGGTACTGAAGTGAAGTACTGTCAAGCAGAGGATGCAGATGCCCGGCTCCCCAATTCC
CACGTTGTATCAGCTCTGAGCCCTGAGCCTGAGGATGGGTAGAGGGGAGAGCCTTCCA
GGTGGACCCCACCTCAGGCAGTGTGACGCTGGGGTCTCCCACCCAGGCAAGGAGAACACA
TCCTGCTCTGGTCTGGCATGGACCTGGCAGGCGAGAGGGTGGCTCAGCAGCACGTGT
GAAGTCGAAGTGCAGTCACAGATATCAATGATCACGCCCTGAGTCATCACTTCCAGAT
TGGGCCTATAAGCCTCCCTGAGGATGTGGAGCCGGACTCTGGTGGCCATGCTAACAGCCA
TTGATGCTGACCTCGAGCCGCCTCCGCCTATGGATTGGCATTGAGAGGGGAGACACA
GAAGGGACTTTGGCCTGGATTGGAGGCCAGACTCTGGCATGTTAGACTCAGACTCTGCAA
GAACCTCAGTTATGAGGCAGCTCCAAGTCAGGAGTGGTGGTGGTGCAGAGTGTGGCGA
AGCTGGTGGGGCCAGGCCAGGCCACGGCACCGCCACGGTACTGTCTAGTGGAGAGA
GTGATGCCACCCCCCAAGTTGGACCAAGGGAGACTACGAGGCCAGTGTCCCCATCAGTGGCC
AGCCGGCTTTCTGCTGACCATCCAGCCCTCCGACCCATCAGCCGAACCTCAGGTTCT
CCCTAGTCATGACTCAGAGGCTGGCTCTGCATTGAGAAATTCTCCGGGGAGGTGCACACC
GCCAGTCCCTGCAGGGCCCCAGCCTGGGACACCTACACGGTCTGTGGAGGCCAGGA
TACAGCCCTGACTCTGGCCCTGTGCCCTCCAAACCTCTGCACACCCGCCAAGACCATG
GCTTGTGATGAGTGGACCCAGCAAGGACCCGATCTGGCCAGTGGCAGGGTCCCTACAGC
TTCACCCCTGGTCCAAACCCACGGTCAACGGGATTGGCGCCTCCAGACTCTCAATGGTT
CCATGCCCTACCTCACCTGGCCCTGCATTGGGTGGAGCCACGTGAACACATAATCCCCGTGG
TGGTCAGCCACAATGCCAGATGTGGCAGCTCTGGTCTGAGTGTGATCGTGTGCGCTGCAAC
GTGGAGGGGCAGTGCATGCGCAAGGTGGCCGATGAAGGGCATGCCACGAAGCTGTGCGG
AGTGGGCATCCTGTAGGCACCCCTGGTAGCAATAGGAATCTTCTCATCCTCATTTCAACCC
ACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCCAGCAGACAGCGTGGCCCTGAAGGCG
ACTGTCTGAATGGCCCAGGCAGCTAGCTGGAGCTGGCCTCTGGCTCCATCTGAGTCCC
CTGGGAGAGAGGCCAGCACCCAAAGATCCAGCAGGGACAGGACAGAGTAGAAGCCCTCCAT
CTGCCCTGGGGTGGAGGCACCATCACCACCAAGGCATGTCTGCAAGGCCCTGGACACCAAC
TTTATGGACTGCCCATGGGAGTGTCCAATGTCAGGGTGTGCTCCAAATAATAAGCCCCA
GAGAAGTGGCTGGGCCATTGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

190/246

FIGURE 188

MVPAWLWLLCVSVPQALPKAQPAELSVEPVENYGGNFPLYLTKLPLPREGAEGQIVLSGDSG
KATEGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHKDENDQVP
HFSQAIYRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQPSPDMFQLEPRL
GALALSPKGSTS LDHALERTYQLLVQVKDMGDQASGHQATATVEVSIESTWVSL
EPIHLAE NLKVLYPHHMAQVHWGGDVHYHLESHPPGPFEVNAEGNLYVTRELDREAQAEYLLQVRAQN
SHGEDYAAPLELHVLMVDENDVPICPPRDPTVSIPELSPPGTEVTRLSAEDADAPGSPNSH
VYVQLLSPEPEDGVEGRAFQVDPTSGSVTLGVPLRAGQNILLV
LAMDLAGAEGGFSSCTCE VEVAVTDINDH
APEFITSQIGPISLPEDVEPGTLVAMLTAIDADLEPAFRLMDFAIERGDTE
GTFGLDWEPD
SGHVRLRLCKNLSYEAA
PSHEVVVVQSVAKLV
GPGPGP
GATATV
TVL
VERV
MPPP
KLDQ
E
SYE
AS
V
P
IS
A
P
A
G
S
F
L
L
T
I
Q
P
S
D
P
I
S
R
T
L
R
F
S
L
V
N
D
S
E
G
W
L
C
I
E
K
F
S
G
E
V
H
T
A
Q
S
L
Q
G
A
Q
P
G
D
T
Y
T
V
L
V
E
A
Q
D
T
A
L
T
I
A
P
V
P
S
Q
Y
L
C
T
P
R
Q
D
H
G
L
I
V
S
G
P
S
K
D
P
D
L
A
S
G
H
G
P
Y
S
F
T
L
G
P
N
P
T
V
Q
R
D
W
R
L
Q
T
L
N
G
S
H
A
Y
L
T
L
A
L
H
W
V
E
P
R
E
H
I
I
P
V
V
V
S
H
N
A
Q
M
W
Q
L
L
V
R
V
I
V
C
R
C
N
V
E
G
Q
C
M
R
K
V
G
R
M
K
G
M
P
T
K
L
S
A
V
G
I
L
V
G
T
L
V
A
I
G
I
F
L
I
L
I
F
T
H
W
T
M
S
R
K
K
D
P
D
Q
P
A
D
S
V
P
L
K
A
T
V

Signal peptide:
amino acids 1-18

Transmembrane domain:
amino acids 762-784

191/246

FIGURE 189

192/246

FIGURE 190

MARMSFVIAACQLVLGLLMTSLTESSIONSECPCQLCVCEIRPWFTPQSTYREATTVDNDLRLTRIPSNLSSDT
QVLLQSNNIAKTVDELQQLFNLTELDFSQNNFTNIKEVGLANLTQLTTLHLEENQITEMTDYCLQDLSNLQEL
YINHNQISTISAHAFAGLKNLLRLHLSNKLKVIDSRWFNSTPNLEILMIGENPVIGILDMMFKPLANLRSIVL
AGMYLTDIPGNALVGLDSLESLSFYDNKLVKVPQLALQKVPNLKFLLDNKNPIHKIQEGDFKNMLRLKELGINN
MGELVSVDRYALDNLPELTKLEATNNPKLSYIHLAFRSVPALESLMLNNNALNAIYQKTVESLPNLREISIHS
NPLRCDCVIHWINSNKTNIRFMEPLSMFCAMPPEYKGHQVKEVLIQDSSEQCLPMISHDSFPNRLNVDIGTTVF
LDCRAMAEPEPEIYWVTPIGNKITVETLSDKYKLLSEGTLIESNIQIEDSGRYTCVAQNVQGADTRVATIKVNG
TLLDGTQVLKIYVKQTESHSILVSWKVNNSVMTSNLKWSATMKIDNPHTITYTARVPVDVHEYNLTHLQPSTDY
EVCLTVSNIHQQTQKSCVNVTTKNAAFAVDISDQETSTALAAVMGSMFAVISLASIAYFAKRFKRKNYHHSLK
KYMQKTSSIPLNELYPPLINLWEGDSEKDGDGSAUTKPTQVDTSRSYMW

Important features:**Signal peptide:**

Amino acids 1-25

Transmembrane domain:

Amino acids 508-530

N-glycosylation sites:Amino acids 69-73; 96-100; 106-110; 117-121; 385-389; 517-521;
582-586; 611-615**Tyrosine kinase phosphorylation site:**

Amino acids 573-582

N-myristoylation sites:

Amino acids 16-22; 224-230; 464-470; 637-643; 698-704

193/246

FIGURE 191

GGGAGAGAGGATAAATAGCAGCGTGGCTTCCCTGGCTCCTCTGCATCCTCCGACCTTC
CCAGCAATATGCATCTTGCACGTCTGGTCGGCTCCTGCTCCCTCCTCTGCTACTGGGGCC
CTGTCTGGATGGCGGCCAGCGATGACCCATTGAGAAGGTATTGAGAAGGGATCAACCGAGG
GCTGAGCAATGCAGAGAGAGAGGTGGCAAGGCCCTGGATGGCATCAACAGTGAATCACGC
ATGCCGGAAGGGAAGTGGAGAAGGTTTCAACGGACTTAGCAACATGGGAGCCACACCAGC
AAGGAGTTGGACAAAGGCCTCAGGGCTCAACCACGGCATGGACAAGGTTGCCATGAGAT
CAACCATGGTATTGGACAAGCAGGAAGGAAGCAGAGAAGCTGGCATGGGTCAACAAACG
CTGCTGGACAGGCCGGGAAGGAAGCAGACAAAGCGGTCCAAGGGTTCCACACTGGGTCCAC
CAGGCTGGGAAGGAAGCAGAGAAACTTGGCAAGGGTCAACCATGCTGCTGACCAGGCTGG
AAAGGAAGTGGAGAAGCTTGGCAAGGTGCCACCATGCTGCTGGCCAGGCCGGGAAGGAGC
TGAGAATGCTATAATGGGTCAACCAAGCCAGCAAGGAGGCCAACAGCTGCTGAATGGC
AACCATCAAAGCGGATCTTCCAGCCATCAAGGAGGGCCACAACCACGCCGTTAGCCTCTGG
GCCCTCAGTCAACACGCCCTTCATCAACCTTCCGCCCTGTGGAGGAGCGTCGCCAACATCA
TGCCCTTAAACTGGCATCCGGCTTGCTGGAGAATAATGTCGCCGTTGTCACATCAGCTGAC
ATGACCTGGAGGGTTGGGGTGGGGACAGGTTCTGAAATCCCTGAAGGGGTTGTACTGG
GATTGTGAATAACCTGATACACCA

194/246

FIGURE 192

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66675
><subunit 1 of 1, 247 aa, 1 stop
><MW: 25335, pI: 7.00, NX(S/T): 0
MHLARLVGSCSLLLLL GALSGWAASDDPIEKVIEGINRGLSNAEREVGKALDGINSGITAG
REVEKVFNGLSNMGSHTGKELDKGVQGLNHGMDKVAHEINHGIGQAGKEAEKLGHGVNNAAG
QAGKEADKAVQGFHTGVHQAGKEAEKLGQGVNHAADQAGKEVEKLGQGAHHAAGQAGKELQN
AHNGVNQASKEANQLLNGNHQSGSSSHQGGATTPLASGASVNTPFINLPALWRSVANIMP
```

Important features of the protein:

Signal peptide:

amino acids 1-25

Homologous region to circumsporozoite (CS) repeats:

amino acids 35-225

195/246

FIGURE 193

GAAGTAGAGGTGTTGTGCTGAGCGGCCGCTCGCGAACACTGTGTGGACCGTCTGCTGGGACTCC
GCCCTGCGTCCGCTCAGCCCCGTGGCCCCGCGCACCTACTGCCATGGGAGACGCGGCCTCGT
CTCGGGGCCACCTGTTGCTGGGCTTCAGTTCCCTGCTCCTCGTCATCTCTCTGATGGACA
TAATGGGCTTGGAAAGGGTTTGGAGATCATATTCAATTGGAGGACACTGGAAGATGGGAAGA
AAGAAGCAGCTGCCAGTGGACTGCCCTGATGGTATTATTCAAAATCCTGGTGTGGAGCT
TCAAAGCTCTAAAGCCAAATTGCAAGATCTACGGAAATTTCAGAACTCTCCATAATT
TGTATGGTAAATCTGAGGATGAAGAGGAACCCAAAGATGAAGATTTCAGCCCTGACGGGG
GTATATTCCACGAATCCTTTCTGGATCCCAGTGGCAAGGTGCATCCTGAAATCATCAAT
GAGAATGGAAACCCAGCTACAAGTATTTCAGTGCAGTGGCAGCAAGTTGTTCAGGGAT
GAAGGAAGCTCAGGAAAGGCTGACGGGTGATGCCTTCAGAAAGAAACATCTGAAGATGAAT
TGTAACATGAATGTGCCCTTCTTCATCAGAGTTAGTGTCTGGAAGGAAAGCAGCAGGGA
AGGAAATATTGAGGAATCATCTAGAACAAATTAGCCGACCAGGAAACCTCATTCTACCTAC
ACTGGAAGGAGCGCTCACTGTGGAAAGAGTTCTGCTAACAGAACGCTGGTCTGCATTTGT
GGATCCAGCGGAGAGTGGCAGACTTCTCCTTCCCTCACCTAAATGTCAACTTGT
CATTGAATGTAAAGAATGAAACCTCTGACACAAAA

196/246

FIGURE 194

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA67300
><subunit 1 of 1, 172 aa, 1 stop
><MW: 19206, pI: 5.36, NX(S/T): 1
METRPRILGATCLLGFSFLLLVISSDGHNGLKGFDHIHWRTLEDGKKEAAASGLPLMVI
IHKSWCGACKALKPKFAESTEISELSHNFVMVNLEDEEEPKDEDFSPDGGYIPRILFLDP
SGKVHPEIINENGNPSYKYFYVSAEQVQQGMKEAQERLTGDAFRKKHLEDEL

Important features of the protein:**Signal peptide:**

Amino acids 1-23

Thioredoxin family proteins:

Amino acids 58-75

N-myristoylation sites:

Amino acids 29-35; 67-73; 150-156

Amidation site:

Amino acid 45-49

197/246

FIGURE 195

CGGCTCGAGTCAGCTGGGGAGATTCAGTGCATTGCCCTGGGTGCTCTTCATCTGGATTTGAAAGT
TGAGAGCAGCAGTGGCTCCACTGAAACTCATCCTGCTGCCAGTGTACTGGATTATCCTGGCCTGAATG
ACTTGAATGTTCCCGCCTGAGCTAACAGTCCATGTGGGTGATTCAAGCTCTGATGGGATGTTCCAGAGC
ACAGAAGACAAATGTATATTCAAGAGTACGGACTCTGTCAACCGAGGACACGCCAAGGACGAATATGTGCT
ATACTATTACTCCAATCTCAGTGTGCTTATGGGCGCTTCCAGAACCGCGTACACTTGATGGGGACATCTTAT
GCAATGATGGCTCTCCTGCTCCAAGATGTGCAAGAGGCTGACCAGGGAACCTATATCTGTGAAATCCGCC
AAAGGGGAGAGCCAGGTGTTCAAGAAGGCGGTGTTACTGCATGTGCTTCAGAGGAGCCAAAGAGCTCATGGT
CCATGTGGGTGGATTGATTCAAGATGGGATGTTCCAGAGCAGAAGTGAACACAGTGAACAGGTAGAAT
GGATATTTCAAGGACGGCGCAGAAGGAGGAGATTGATTTGTTACTACCACAAACTCAGGATGTTGAG
TACTCCCAGAGCTGGGCCACTTCCAGAACCTGTGAACTGGTGGGGACATTTCCGCAATGACGGTCCAT
CATGCTTCAAGGAGTCAGATGGAGGAAACTACACCTGCAGTATCCACCTAGGGAACCTGGTGTCA
AGAAAACCATTGTGCTGCATGTCAGCCCGAAGAGCCTCGAACACTGGTACCCCGGCAGCCCTGAGGCCTCTG
GTCTGGGTGTTAATCAGTTGGTGTGATATTGTTGAAATTGCTGTGCCCCATCTGCTGCTCCCTGTTCTGAT
ATTGATCGTGAAGAAGACCTGTGAAATAAGAGTTCAAGTGAATTCTACAGTCTGGTGAAGAACACGAAGAAGA
CTAATCCAGAGATAAAAGAAAAACCCCTGCCATTGAAAGATGTGAAAGGGAGAACACATTTACTCCCCAATA
ATTGATCGGAGGTGATCCAGGAAGAACCAAGTGAAGGGGAAACTGACCCATGCACCCAGT
TTGGCCTCTCTGAGGTCAAGATCGGAACAACACTCAGTGAAAAAAAGTCAGGTGGGGGAAATGCCAAAAACACAGC
AAGCCTTTTGAGAAGAATGGAGAGTCCTCATCTCAGCAGCGGTGGAGACTCTCTCTGTGTTGCTGGC
CACTCTACCACTGATTCAGACTCCCGCTCTCCAGCTGTCTCCTGTCTCATTGTTGGTCAATACACTGAAG
ATGGAGAATTGGAGGCTGGCAGAGAGACTGGACAGCTGGAGGAACAGGCCTGCTGAGGGAGGGAGCATG
GACTTGGCCTCTGGAGTGGGACACTGCCCTGGGAACCAGGCTGAGCTGAGTGGCCTCAACCCCCCGTTGGAT
CAGACCCCTCTGTGGCAGGGTCTTAGTGGATGAGTTACTGGGAAGAATCAGAGATAAAACCAACCCAAATCAA

198/246

FIGURE 196

MFCPLKLILLPVLLDYSLGLNLDLNVSPPPELTVHVGDSALMGCVFQSTEDKCIFKIDWTLSPEHAKDEYVLYYY
SNLSVPPIGRFQNRVHLMGDILCNDGSLLLQDVQEADQGTYICEIRLKGESQVFKKAVVLHVLPEEPKELMVHVG
GLIQMGCVFQSTEVKHVTKVEWIFSGRRAKEEIVFRYYHKLRLMSVEYSQSWGHFQNRVNLVGDIFRNDGSIMLQ
GVRESDGNNYTCIHLGNLVFKKTIVLHVSPEEPRTLVTPAALRPLVLGGNQLVIIIVGIVCATILLLPVLLIV
KKTGKSSVNSTLVKNTKKTNPETKEKPCHFERENCEGEKHIYSPIIVREVIEEEEPSEKSEATYMTMHPVWPS
LRSDRNNNSLEKKSGGGMPKTQQAF

199/246

FIGURE 197

CGCCATGGCCGGGCTATCCCGCGGGTCCGCGCGCACTGCTCGCCGCCCTGCTGGCGTCGACG
CTGTTGGCGCTGCTCGTGTGCGCCCGCGCGGGGTGCGCGGCCGGGACCACGGGGACTGGGA
CGAGGCCTCCCGGCTGCCGCCGCTACCAACCCCGCGAGGACGCGGCCGCGTGGCCCGCTTCG
TGACGCACGTCTCCGACTGGGGCGCTCTGGCCACCCATCTCACGCTGGAGGCGGTGCGCGGC
CGGCCCTTCGCCGACGTCTCGCTCAGCGACGGGCCCCCGGCCGCGGGCAGCGCGTGCC
CTATTCTACCTGAGCCCGCTGCAGCTCTCGTGAGCAACCTGCAGGAGAATCCATATGCTA
CACTGACCATGACTTGGCACAGACCAACTCTGCAAGAAACATGGATTGATCCACAAAGT
CCCCTTGTGTTCACATAATGCTGTCAGGAACGTGACCAAGGTGAATGAAACAGAAATGGA
TATTGCAAAGCATTGTTATTGACACCCCTGAGATGAAAACCTGGCCTTCAGCCATA
ATTGGTTCTTGCTAAGTTGAATATAACCAATATCTGGGTCTGGACTACTTGGTGGACCA
AAAATCGTGACACCAGAAGAATTATAATGTCACAGTTCATGAAGCAGACTGTGGTGAAT
TTAGCAACACTTATGAAGTTCTAAAGTGGCTACACACTTAAAGGCTTAATGTTCT
CTGGAAAGCGTCCCAGAATATTAGCCAGTTCTGTC

200/246

FIGURE 198

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71269
><subunit 1 of 1, 220 aa, 1 stop
><MW: 24075, pI: 7.67, NX(S/T): 3
MAGLSRGSRALLAALLASTLLALLVSPARGRRGDHGDWDEASRLPPLPPREDAARVAR
FVTHVSDWGALATISTLEAVRGRPFADVLSLSDGPPGAGSGVPYFYLSPLQLSVSNLQEN
PYATLTMTLAQTNFCKKHGFDQSPLCVHIMLSGTVTKVNETEMDIAKHSLFIRHPEMKT
WPSSHNWFFAKLNITNIWVLDYFGGPKIVTPEEYYNVTVQ
```

Important features of the protein:**Transmembrane domain:**

Amino acids

11-29

N-glycosylation sites:

Amino acids

160-164;193-197;216-220

N-myristoylation sites:

Amino acids

3-9;7-13;69-75;97-103

201/246

FIGURE 199

TCGCC**ATGGC**CTCTGCCGGAATGCAGATCCTGGGAGTCGTCTGACACTGCTGGCTGGGTG
AATGGCCTGGTCTCCTGTGCCCTGCCATGTGGAAAGGTGACCGCTTCATCGGCAACAGCAT
CGTGGTGGCCCAGGTGGTGTGGGAGGGCTGTGGATGTCCTGCGTGGTGCAGAGCACCGGCC
AGATGCAGTGCAGGTGTACGACTCACTGCTGGCGCTGCCACAGGACCTGCAGGCTGCACGT
GCCCTCTGTGTCATGCCCTCCTGTGGCCCTGTTGGCTTGCTGGTCTACCTTGCTGGGC
CAAGTGTACCACTGTGTGGAGGAGAAGGATTCCAAGGCCCCCTGGTGCCTCACCTCTGGGA
TTGTCTTGTCATCTCAGGGGTCTGACGCTAATCCCCGTGTGCTGGACGGCGATGCCATC
ATCCGGGACTTCTATAACCCCCCTGGTGGCTGAGGCCAAAAGCAGGGAGCTGGGGCCTCCCT
CTACTTGGGCTGGGCCCTCAGGCCTTTGTTGCTGGTGGGGTTGCTGTGCTGCACCT
GCCCTCGGGGGGGTCCCAGGGCCCCAGCCATTACATGGCCCGCTACTCAACATCTGCCCT
GCCATCTCTCGGGGGCCCTCTGAGTACCCCTACCAAGAATTACGTC**TGAC**GTGGAGGGGAATG
GGGGCTCCGCTGGCGTAGAGCCATCCAGAAGTGGCAGTGCCAACAGCTTGGATGGTT
CGTACCTTTGTTCTGCCCTGCTATTTCCTTGACTGAGGATATTAAATTCAATT
GAAAAGTGGAGCCAAAGAGGGGATGCTTGAGATTCTGGATCTGACATGCCATCTAGAAC
CAGTCAAGCTATGGAACATAATGCCAGGCTGCTGCTGTGCTGGCTTGCAACAAGACAGAC
TGTCCCCAAGAGTTCTGCTGCTGGCTGGGGCTGGCTCCCTAGATGTCAGTGGACAGCTG
CCCCCCATCCTACTCAGGTCTCTGGAGCTCCTCTTCAACCCCTGGAAAAAAACAAATCATCTG
TTAACAAAGGACTGCCACCTCCGGAACCTCTGACCTCTGTTCCCTCCGTCTGATAAGACG
TCCACCCCCCAGGGCCAGGTCCCAGCTATGTAGACCCCCGCCACCTCCAACACTGCACC
CTTCTGCCCTGCCCTCGTCTCACCCCTTACACTCACATTATCAAATAAGCATG
TTTGTAGTGCATGCA

202/246

FIGURE 200

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73736
><subunit 1 of 1, 220 aa, 1 stop
><MW: 23292, pI: 8.43, NX(S/T): 0
MASAGMQILGVVLTLGVNGLVSCALPMWKTAFIGNSIVVAQVVWEGLWMSCVVQSTGQM
QCKVYDSLLALPQDLQAARALCVIALLVALFGLLVYLAGAKCTTCVEEKDSKARLVLTSGIV
FVISGVLTLLIPVCWTAAHAIIRDFYNPLVAEAQKRELGASLYLGWAASGLLLGGGLLCCTCP
SGGSQGPGSHYMARYSTSAPAISRGPSEYPTKNYV
```

Transmembrane domains:

amino acids 8-30 (type II), 82-102, 121-140, 166-186

203/246

FIGURE 201

AGTGACAATCTCAGAGCAGCTTCTACACCACAGCCATTCCAGC**ATGAAGATCACTGGGGTCTCCTCTGCTC**
TGTACAGTGGCTATTTCTGTAGCAGCTCAGAAGCTGCTAGTCTGCTCAAAAAAGTGGACTGCAGCATT
CAAGAAGTATCCAGTGGGGCCATCCCCTGCCCATCACATACCTACCAAGTTGTGGTTCTGACTACATCACCT
ATGGAATGAATGTCACTGTGTACCGAGAGCTGAAAAGTAATGGAAGAGTTCACTTCTTCACGATGGAAGT
TGCTAAATTCTCCATGGACATAGAGAGAAAGGAATGATATTCTCATCATCTTCATCATCCCAGGCTCTGAC
TGAGTTCTTCAGTTTACTGATGTTCTGGGTGGGGACAGAGCCAGATTCAAGAGTAATCTTGAATGGA
GAAAGTTCTGTGCTACCCCTACAAACCCATGCCTCACTGACAGACCAGCATTTTTTTAACACGTCAATAA
AAAAATAATCTCCCAGA

204/246

FIGURE 202

MKITGGLLLLCTVVYFCSSSEAASLSPKKVDCSIYKKYPVVAIIPCPITYLPVCGSDYITYGNECHLCTESIKN
GRVQFLHDGSC

Important features:

Signal peptide:

amino acids 1-19

205/246

FIGURE 203

CGACGATGCTACGCGGCCGGCTGCCCTCCGGACCTCCGTAGCGCTGCCGCCCTGGCTGCCGCTG
CTCTCGTCGCTTGCAGCTGCTCTCTAGAGCGAGGGACCCGGTGGCCTCGTCGCTCAGCCCTATTCGG
CACCAAGACTCGCTACGAGGATGTCAACCCCGTGCATTGTCGGGCCCGAGGCTCCTCGTGGCGGGACCCCTGAGC
TGGTGGAGGGGACCTGCACCCCGTGCAGCTGGTCGCCCTCATTCGCCACGGCACCCGCTACCCACGGTCAAA
CAGATCCGCAAGCTGAGGCAGCTGCACGGGTTGCAGGGCCGGTCCAGGGATGGCGGGCTAGTAGTAC
CGGCAGCCCGACCTGGGTGCAGCGCTGGCCGACTGGCTTGTGGTACGCGGACTGGATGGACGGGAGCTAG
TAGAGAAGGGACGGCAGGATGCGACAGCTGGCGCTGCCTCGCTTCCCGGCCCTTTCAGCCGT
GAGAACTACGCCGCTCGGCTCATCACCAGTCCAAGCACCGCTGCATGGATAGCAGCGCCGCTTCGCA
GGGGCTGTGGCAGCACTACCACCTGGCTGCCGCCGACGTGCGAGATATGGAGTTGGACCTCCAACAG
TTAATGATAAAACTAATGAGATTTTGATCACTGTGAGAAGTTTTAATGAGTAGAAAAAAATGCTACAGCT
CTTATCACGTGGAAGCCTCAAAACTGGACCAGAAATGCAAGAACATTAAAAAGTTGCAAGCTACTTGCA
AGTGCAGTAAATGATTAAATGCAAGATTTAATTCAAGTAGCCTTTCACCTGTTATTGACCTGGCAATT
AAGGTGTTAAATCTCTGGTGTGATGTTTGACATAGATGATGCAAAAGGTATTAGAATATTAAATGATCTG
AAACAATATTGAAAAGAGGGATATGGGTATACTATTAAACAGTCGATCCAGCTGCACCTGTTAGGATATT
TCAGCACTGGACAAAGCAGTGAAGAACAAAGGTCTCAGCCAATTCTCTCAGTCATCCCTCAGTTG
GTCATGCAGAGACTCTTCTTCACTGCTTCTCATGGGCTACTTCAGAACAGAACGGAAACCCCTAACAGCGTAC
AATTACAAAAAAACAAATGCATCGGAAGTCCGAAGTGGTCTCATGGTACCTTATGCCCTCGAACCTGATATT
GCTTACACTGTGAAAATGCTAAAGACTCTAAAGAACATCCGAGTGCAGATGTATTAAATGAAAAGGTGT
TACCTTGGCTTACTCACAAAGAAACTGTTTCAATTGAGATCTGAAGAACCACTAACAGGACATCCTCAG
AGTTGTCAAACCAGTGAAGAATGTGAATTAGCAAGGGTAACAGTACATCTGATGAACATGAGTAACTGAAGA
ACATTAAATTCTTAGAATCTGCAATGAGTGAATTACATGCTTGTAAAGTAGGAGCAATTCTTGATTACAG
GAAGCTTTATATTACTTGAGTATTCTGTCTTTCACAGAAAACATTGGGTTCTCTGGGTTGGACATG
AAATGTAAGAAAAGATTTCACTGGAGCAGCTCTTAAGGAGAACAAATCTATTAGAGAACAGCTGCC
CTGCAAATTTACAGAAATGAAATTCTCCTACTTATATAAGAAATCTCACACTGAGATAGAATTGTGATTTC
ATAATAACACTTGAAAAGTGGAGTAACAAATATCTCAGTTGGACATCCTTAACCTGATTGAACTGCTCA
GGAACCTTACAGATTGTTCTGCAGTTCTCTCTTCCCTCAGGTAGGACAGCTTAGCATTCTTAATCAG
GAATATTGTGTAAGCTGGGAGTACTCTGGAAAGTAACATCTCCAGATGAGAACATTGAAACAAGAAC
AGAGTGTGAAAAGGACACCTCACTGAAGCAAGTCGGAAAGTACAATGAAAATAATATTGTTGGTATT
TTATGAAATATTGAAACATTTCATAATTCTTTACTCTAGGAAGTCTCAAAAGACCATCTTAAATTA
TTATGTTGGACAATTAGCAACAGTCAAGATGTTAGAATCGAAGGTTTCAAATCCATTGCTTAGCTA
TTTCAATTGTCACCTGGCTCGATTATATTCTCTTCTATTATGAAATGTTGCTTGGTGTGATT
TTCTTCTTCTTGTAAATGTTGAGTTGTCACCTGGCTGAAAGTATTGCTATAATAAGAAAATTC
TTGTGACTTAAAAAA

206/246

FIGURE 204

MLRAPGCLLRTSVAPAAALAAALLSSLARCSLLEPRDPVASSLSPYFGTKTRYEDVNPVLLSGPEAPWRDPELL
EGTCTPVQLVALIRHGTRYPTVKQIRKLRLQHGLLQARGSRDGGASSTGSRDLGAALADWPLWYADWMDGQLVE
KGQDMRQLALRLASLFPALFSRENRYGRLRLITSSKHRCMDSSAAFLQGLWQHYHPLGPPPDVADMEFGPPTVN
DKLMRFFDHCEKFLTEVEKNATALYHVEAFKTGPEMQNILKKVAATLQVPVNLDLNADLIQVAFFTCSFDLAIKG
VKSPWCDVFDIDDAKVLEYLNDLKQYWKRGYGYTINSRSSCTLFQDIFQHLDKAVEQKORSQPISSPVILQFGH
AETLLPLLSIMGYFKDKEPLTAYNYKKQMHRKFRSGLIVPYASNLIIFVLYHCENAKTPKEQFRVQMILLNEKVLP
LAYSQETVSFYEDLKNHYKDILQSCQTSEECELARANSTSDEL

Important features:

Signal sequence

amino acids 1-30

N-glycosylation sites.

amino acids 242-246, 481-485

N-myristoylation sites.

amino acids 107-113, 113-119, 117-123, 118-124, 128-134

Endoplasmic reticulum targeting sequence.

amino acids 484-489

207/246

FIGURE 205

GGCAGCGCGCGGGCGGCGAGAGGAAACGCGGCCGGGCCGGCCCTGGAGA**TG**
GTCGGCGCGCCGCGGGCTGGTGTCTCGTGTCTGGCTCCCCGCGTGCCTCGCGGCCA
CGGCTTCCGTATCCATGATTATTGTACTTCAGTGCTGAGTCCTGGGACATTGATACA
TCTTCACAGCCACACCTGCCAAGGACTTGGTGGTATCTTCACACAAGGTATGAGCAGATT
CACCTTGTCGGCGTGAACCTCCAGAGGCCCTGCCGGGAACTCAGCAACGGTTCTCATCCA
GGACCAAGATTGCTCTGGTGGAGAGGGGGCTGCTCCTCTCCAAGACTCGGGTGGTCC
AGGAGCACGGCGGGCGGGCGGTGATCATCTTGACAACGCAGTTGACAATGACAGCTTCTAC
GTGGAGATGATCCAGGACAGTACCCAGCGCACAGCTGACATCCCCGCCCTTCCCTGCTCGG
CCGAGACGGCTACATGATCCGGCGCTCTGGAACAGCATGGGCTGCCATGGGCCATCATTT
CCATCCCAGTCATGTCACCAGCATCCCCACCTTGAGCTGCTGCAACCGCCCTGGACCTTC
TGG**TAGAAGAG**TTGCTTGTCCCACATTCCAGCCATAAGTGACTCTGAGCTGGAAAGGGAAACCC
AGGAATTGCTACTTGAATTGGAGATAGCATCTGGGACAAGTGAGCCAGGTAGAGGA
AAAGGGTTGGCGTTGCTAGGCTGAAAGGGAAGGCCACACCCTGGCCTCCCTCCCCAGG
GCCCCAAGGGTGTCTCATGCTACAAGAAGAGGCAAGAGACAGGCCAGGGCTTGTGCTA
GAACCCGAAACAAAAGGAGCTGAAGGCAGGTGGCCTGAGAGGCCATCTGTGACCTGTACACT
CACCTGGCTCCAGCCTCCCTACCCAGGGTCTCTGCACAGTGACCTTCACAGCAGTTGTTGG
AGTGGTTAAAGAGCTGGTGTGTTGGGACTCAATAAAACCTCACTGACTTTAGCAATAAA
GCTTCTCATCAGGGTGCAAAAAAAAAAAAAAA

208/246

FIGURE 206

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76532
><subunit 1 of 1, 188 aa, 1 stop
><MW: 21042, pI: 5.36, NX(S/T): 2
MVPGAAGWCLVLWLPACVAAHGFRIGHDLYFQVLSPGDIRYIFTATPAKDFGGIFHTRYEQ
IHLVPAEPPEACGELSNGFFIQDQIALVERGGCSFLSKTRVVQEHGGRAVIISDNAVDNSF
YVEMIQDSTQRTADIPALFLLGRDGYMIRRSLEQHGLPWAIISIPVNVTSIPTFELLQPPWTFW
```

Signal peptide:
amino acids 1-20

209/246

FIGURE 207

CTCGCTTCTTCCTTCTGGATGGGGGCCAGGGGCCAGGAGAGTATAAAGGCATGTGGAG
GGTGCAGGCACAACCAGACGCCAGTCACAGGCGAGAGCCTGGG**ATG**CACCGGCCAGAGG
CCATGCTGCTGCTGCTCACGCTTGCCTCCTGGGGGCCACCTGGCAGGGAAAGATGTAT
GGCCCTGGAGGAGGCAAGTATTCAGCACCCTGAAGACTACGACCATGAAATCACAGGGCT
GCGGGTGTCTGTAGGTCTTCTCCTGGTGAAGACTGGAGACTCCTGGG
ACGTGAAACTGGGAGCCTAGGTGGGAATACCCAGGAAGTCACCCTGCAGCCAGGCGAATAC
ATCACAAAAGTCTTGTGCCTCCAAGCTTCCTCCGGGTATGGTCATGTACACCAGCAA
GGACCGCTATTCTATTGGAGCTTGATGGCCAGATCTCCTCTGCCTACCCAGCCAAG
AGGGGCAGGTGCTGGTGGCATCTATGCCAGTATCAACTCCTGGCATCAAGAGCATTGGC
TTTGAATGGAATTATCCACTAGAGGAGCCGACCACTGAGCCACCAAGTTAATCTCACATACTC
AGCAAACTCACCGTGGCTCGC**TAGGGTGGGTATGGGCCATCGAGCTGAGGCCATCTGT**
GTGGTGGTGGCTGATGGTACTGGAGTACTGAGTCGGGACGCTGAATCTGAATCCACCAATA
AATAAAGCTCTGCAGAAAA

210/246

FIGURE 208

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76541
><subunit 1 of 1, 178 aa, 1 stop
><MW: 19600, pI: 5.89, NX(S/T): 1
MHRPEAMLLLTLALLGGPTWAGKMYGPGGGKYFSTTEDYDHEITGLRVSGLLLVKSVQVK
LGDSDVDVKLGALGGNTQEVTLQPGEYITKVFVAFQAFLRGMVMYTSKDRYFYFGKLDGQISS
AYPSQEGQVLVGIYGQYQLLGIKSIGFEWNYPLEEPTTEPPVNLTYSANSPVGR
```

Signal peptide:
amino acids 1-22

211/246

FIGURE 209

212/246

FIGURE 210

MERAVRVESGVLVGVVCLLLACPATATGPEVAQPEVDTTLGRVRGRQGVKGTDRLNVFLGIPFAQPPLGPDR
FSAPHQPAQPWEGVRDASTAPPMCLQDVESTMNSSRFVLNGKQQIFSVDCLVLNVYSPAEPAGSGRPVMVWH
GGALITGAATSYDGSALAAYGDVVVTVQYRLGVLGFFFSTGDEHAPGNQGFLDVVAALRWQENIAPFGGDLNC
VTVFGGSAGGSIIISGLVILSPVAAGLFHRAITQSGVITTPGIIDSHPWLQKIANTLACSSSPAEMVQCLQQK
EGEELVLSKKLKNTIYPLTVGDGTVPKSPKELLKEKPFHSVPFLMGVNNEFSWLIPRGWGLLDTMEQMSREDM
LAISTPVLTSLDVPPPEMMPTVIDEYILGSNSDAQKCQAFQEFMGDVFINVPTVSFSRYLRSQSPVFFYEFQHR
PSSFAKIKPAWVKADHGAEGAFVFGGFLMDESSRLAFPEATEEEKQLSLTMMAQWTHFARTGDPNSKALPPWP
QFNQAEQYLEINPVPRAGQKFREAWMQFWSETLPSKIQQQWHQKQKNRKAQEDL

Important features:

Signal peptide:
amino acids 1-27

Transmembrane domain:
amino acids 226-245

N-glycosylation site.
amino acids 105-109

N-myristylation sites.
amino acids 10-16, 49-55, 62-68, 86-92, 150-156, 155-161, 162-168, 217-223,
227-233, 228-234, 232-238, 262-268, 357-363, 461-467

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 12-23

Carboxylesterases type-B serine active site.
amino acids 216-232

213/246

FIGURE 211

AACTTCTACATGGGCCTCCTGCTGCTGGTCTTCCTCAGCCTCTGCCGGTGGCCTACAC
CATCATGTCCCTCCACCCCTCTTGACTGCGGGCCGTTAGGTGCAGAGTCTCAGTTGCC
GGGAGCACCTCCCTCCGAGGCAGTCTGCTCAGAGGCCTCGGCCAGAATTCCAGTTCTG
GTTTCATGCCAGCCTGTAAAAGGCCATGGAACCTTGGGTGAATCACCGATGCCATTAAAGAG
GGTTTCTGCCAGGATGGAATGTTAGGTCGTTCTGTCGCTGCTGTTATTCACTAGGCC
ACCAGCCACCTGTGGCGTTGAGTGCTGAAATGAAACTGAGAAAATTAAATTCTCATGT
ATTTTCTCATTATTATTAAATTAAACTGATAGTTGTACATATTGGGGTACATGTGA
TATTGGATACTGTATAACAATATAATGATCAAATCAGGGTAACGGGATATCCATCACA
TCAAACATTATTTTATTCTTTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGC
AGTGGTGCCATCTCAGCTACTGCAACCTCTGCCAGGTTCAAGCGATTCTCATGCC
CACCTCCCAAGTAGCTGGGACTACAGGCATGCACCAATGCCCAACTAATTTGTATT
TAGTAGAGACGGGTTTGCATGTTGCCAGGCTGGCCTGAACTCTGGCCTCAAACAA
CCACTTGCCCTGGCCTCCAAAGTGTATGATTACAGGCGTGAGCCACCGTGCCTGGC
ACATTATCTTCTTGTTGGAACTTGAAATTATAACAATGAATTATTGTTAACTGTC
ATCTCCCTGCTGTGCTATGGAACACTGGGACTTCATCCCCACTCCTCTATC
CAGTTAACCAACCGTACTTCATCCCCACTCCTCTATCCTCCAAACCTCTGATCAC
TTCTACTCTACCTCCATGAGATCCACTTTAGCTCCACATGTGAGTAAGAAAATGCA
ATATTGTCTTCTGTCCTGGCTTATTCACTAACATAATGACTTCTGTTCC
TTGCTGCAAATGACAGGATTCGTTCTTAATTCAATTAAACCACACATGGCA
AAAAA

214/246

FIGURE 212

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPGRGSLLRGPRPRIPVLVSC
QPVKGHGTLGESPMPPKRVFCQDGNNVRSFCVCAVFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

215/246

FIGURE 213

AGGGCCCGGGGTGGAGAGAGCGACGCCGAGGGGATGGCGGCAGCGTCCGGAGCGCCTCT
GGCTGGCGCTACTGCTGCTGGCACTTGGCAGCAGCGCGCCGCTCCGGCGTCTT
CCAGCTGCAGCTGCAGGAGTTCAACGAGCGGGCGTACTGCCAGTGGCGGCGTCTGCG
AGCCCGGCTGCCGGACTTCTCCGCGTCTGCCCTAACGACTTCAAGGCCAGTGGCGGCGTCTCGCC
GGACCCCTGCACCTCAGGACCGTCTCCACGCCGGTATTGGCACCACGCCCTGCCTGCGTCCG
GGACGACAGTAGCGGGGGGGCGCAACCTCTCAACTGCCCTCAATTACCTGGCCGG
GTACCTCTCGCTCATCATCGAAGCTTGGCACGCCAGGAGACGACCTGCCAGAGGCC
TTGCCACCAGATGCACTCATCAGCAAGATGCCATCCAGGCTCCCTAGCTGTGGGTAGAA
CTGGTTATTGGATGAGCAAACCAAGCACCCCTCACAGGCTGCCTACTCTTACGGGTCT
GCAGTGACAACACTATGGAGACAACACTGCTCCCGCTGTGCAAGAAGCGCAATGACCACTTC
GCCACTATGTGTGCCAGCCAGATGGCAACTTGTCTGCCCTGCCGGTGGACTGGGAATA
TTGCCAACAGCCTATCTGTCTTCAGGCTGTATGAAACAGAATGGCTACTGCAGCAAGCCAG
CAGAGTCGCTCTGCCGCCAGGCTGGCAGGGCCGGCTGTGTAACGAATGCATCCCCAACAAAT
GGCTGTCGCCACGGCACCTGCAAGCACTCCCTGGCAATGTACTTGTGATGAGGGCTGGGGAGG
CCTGTTTGTGACCAAGATCTCAACTACTGCACCCACACTCCCCATGCAAGAATGGGGCAA
CGTGCTCCAACAGTGGCAGCGAAGCTACACCTGCACCTGCGCCAGGCTACACTGGTGTG
GACTGTGAGCTGGAGCTCAGCGAGTGTGACAGCAACCCCTGCGCAATGGAGGAGCTGTAA
GGACCAAGGAGGATGGCTACCACCTGCGCTGTGCTCCGGCTACTATGCCCTGCACTGTGAAAC
ACAGCACCTTGAGCTGCGCGACTCCCCCTGCTTAATGGGGCTCTGCCGGAGCGCAAC
CAGGGGCCAACTATGCTGTGAATGTCCCCCAACTTCACCGGCTCCAACGCGAGAAGAA
AGTGGACAGGTGCAACAGCAACCCCTGTGCAACGGGGACAGTGCCTGAACCGAGGTCAA
GCCGCATGTGCCGCTGCCGTCTGGATTCAACGGGCACCTACTGTGAACTCCACGTCAGCGAC
TGTGCCGTAACCTTGCGCCACGGTGGCACTTGCCTGACCTGGAGAATGGGCTATGTG
CACCTGCCCTGCCGGCTCTGGCGACGCTGTGAGGTGCGGACATCCATGATGCCGTG
CCTCGAGTCCCTGCTCAACAGGGCACCTGCTACACCGACCTCTCCACAGACACCTTGTG
TGCAACTGCCCTATGGCTTGTGGCAGCCGCTGCGAGTTCCCGTGGCTGCCGCCAG
CTTCCCTGGTGGCGTCTCGCTGGGTGTGGGCTGGCAGTGTGCTGGTACTGCTGGCA
TGGTGGCAGTGGCTGTGCGGAGCTGCGCTTCGACGGCCGGACAGGGCAGCAGGGAGCC
ATGAACAACCTGTCGGACTTCCAGAAGGACAACCTGATTCTGCCGCCAGCTAAAAACAC
AAACCAAGAAGGAGCTGGAGTGGACTGTGGCTGGACAAGTCCAACGTGGCAAACAGC
AAAACCACACATTGGACTATAATCTGGCCCCAGGGCCCTGGGCGGGGGACCATGCCAGGA
AAGTTCCCCACAGTGACAGAGCTTAGGAGAGAAGGCAGCCACTGCCGTACACAGTGA
GCCAGAGTGCGGATATCAGCGATATGCTCCCCCAGGGACTCCATGTA
GGACATCCCTGCTCAGCCCCGGCTGGACCTTCCCTGCATTGTTACA
TGATATCAGAGGAGAGGAATGAATGTGTCATTGCCACGGAGGTATGGCAGGAGCCTACCT

216/246

FIGURE 214

MAAASRSASGWALLLVALWQQRAAGSGVQLQLQEFINERGVLASGRPCEPGCRTFFRVCL
KHFQAVVSPGPCTFGTVSTPVLGTONFAVRDDSSGGGRNPLQLPFNFTWPGTFSLIIEAWA
PGDDLRLPEALPPDALISKIAIQSLAVGQNWLDEQTSTLRLRYSYRVCSDNYYGDNCSR
LCKKRNDHFGHYVCQPDGNLSCLPGWTGEYCQQPICLSCGHEQNGYCSKPAECLCRPGWQGR
LCNECIPHNGCRHGTCTPWQTCDEGWGGLFCQDLYNCTHSPCKNGATCSNSGQRSYTC
TCRPGYTGVDCELELSECDSPCRNGGCKDQEDGYHCLCPPGYYGLHCEHSTLSCADSPCF
NGGSCRERNQGANYACECPPNFTGSNCEKKVDRCTSNSPANCQCLNRGSPRMCRPGFTG
TYCELHVSDCARNPCAHGGTCHDLENGLMCTCPAGFSGRRCEVRTSIDACASSPCFN RATCY
TDLSTDTFVCNCPYGFVGSRCEFPVGLPPSF PWAVSLGVGLAVLLVLLGMVA VAVRQLRLR
RPDDGSREAMNNLSDFQKDNLIPAAQLKNTNQKKELEVDCGLDKSNCGKQQNHTLDYLNLAGP
PLGRGTMPGKFPHSDKSLGEKAPRLHSEKPECRISAICSPRDSMYQSVCLISEERNECVIA
TEV

Important features of the protein:**Signal peptide:**

amino acids 1-26

Transmembrane domain:

amino acids 530-552

N-glycosylation sites.

amino acids 108-112, 183-187, 205-209, 393-397, 570-574, 610-614

Glycosaminoglycan attachment site.

amino acids 96-100

Tyrosine kinase phosphorylation site.

amino acids 340-347

N-myristoylation sites.amino acids 42-48, 204-210, 258-264, 277-283, 297-303, 383-389,
415-421, 461-467, 522-528, 535-541, 563-569, 599-605, 625-631**Amidation site.**

amino acids 471-475

Aspartic acid and asparagine hydroxylation site.

amino acids 339-351

EGF-like domain cysteine pattern signature.amino acids 173-185, 206-218, 239-251, 270-282, 310-322,
348-360, 388-400, 426-438, 464-476, 506-518**Calcium-binding EGF-like:**amino acids 224-245, 255-276, 295-316, 333-354, 373-394,
411-432, 449-470

217/246

FIGURE 215

CGCGAGGCGCGGGGAGCCTGGGACCAGGAGCGAGAGCCGCCTACCTGCAGCCGCCACG
GCACGGCAGGCCACCATGGCGCTCCTGCTGTGCTTCGCTCCTGTGCGGAGTAGTGGATTTC
GCCAGAAGTTTGAGTATCACTACTCCTGAAGAGATGATTGAAAAAGC_{AAAGGGAAACTGC}
CTATCTGCCATGCAAATTACGCTTAGTCCCAGACCGAGGGACGCTGGACATCGAGTGGC
TGATATCACCAGCTGATAATCAGAAGGTGGATCAAGTGATTATTTATATTCTGGAGACAAA
ATTATGATGACTACTATCCAGATCTGAAAGGCCAGTACATTACGAGTAATGATCTCAA
ATCTGGTGTGATGCATCAAATAATGTAACGAATTACAACGTGTCAGATATTGGCACATATCAGT
GCAAAGT_{AAAAAGCTCCTGGTGTGCAAATAAGAAGATT}CATCTGGTAGTTCTGTTAAG
CCTCAGGTGCGAGATGTTACGTTGATGGATCTGAAGAAATTGGAAGTGACTTTAAGATAAA
ATGTGAACCAAAAGAAGGTTCACTCCATTACAGTATGAGTGGC_{AAAAATTGTCGACTCAC}
AGAAAATGCCCACTTCATGGTTAGCAGAAATGACTTCATCTGTTATATCTG_{AAAAATGCC}
TCCTCTGAGTACTCTGGGACATACAGCTGTCAGTCAGAACAGAGTGGCTCTGATCAGTG
CCTGTTGCGTCTAACGTTGCTCCCTCCTCAAATAAGCTGGACTAATTG_{CAGGAGCCATTA}
TAGGAAC_{TTGCTCTAGCGCTCATGGCTTATC}ATCTTTGCTGTCG_{AAAAAGCGC}
AGAGAAGAAAAATATGAAAGAAGTTCACTCACGATATCAGGGAAAGTGTGCCACCTCCAAA
GAGCCGTACGTCCACTGCCAGAGCTACATCGGCA_{GTAATCATT}CATCCCTGGGTCCATGT
CTCCTTCCAACATGAAAGGATATTCCAAGACTCAGTATAACCAAGTACCAAGTGAAGACTTT
GAACGCACTCCTCAGAGTCCGACTCTCCACCTGCTAAGTTCAAGTACCC_{CTACAAGACTGA}
TGGAATTACAGTTGTATAAATATGGACTACTGAAGAATCTGAAGTATTGTATTATGACTT
TATTTAGGCCTCTAGTAAAGACTTAAATGTTTTAAAAAAAGCACAAGGCACAGAGATTA
GAGCAGCTGTAAGAACACATCTACTTTATGCAATGGCATTAGACATGTAAGTCAGATGT_{CAT}
GTC_{AAAATTAGTACGAGCCAAATTCTTGT}AAAAAACCTATGTATAGT_{GACACTGATAGT}
TAAAAGATGTTTATATATTTC_{AA}ACTACC_{ACTAACA}ATT_{TT}TAAC_{TT}TCATATGC
ATATTCTGATATGTGGCTTTAGGAAAAGTATGTTAATAGTTGATTTCAAAGGAAATT
TTAAAATTCTTACGTTCTGTTAATGTTTGT_{CATTAGTAA}ATACATTGAAGGGAAATA
CCCGTTCTTCCCTTTATGCACACAACAGAAACACGCGTTGT_{CATGCCTCAA}ACTATT
TTTATTGCAACTACATGATTCACACAATTCTTAAACAACGACATAAAATAGATTCT
TGTATATAAAACTACATACGCTCCATAAGTA_{AAATTCTCAAAGGTGCTAGAACAAATCG}
TCCACTTCTACAGTGTCTCGTATCCAACAGAGTTGATGCACAATATATAAAACTCAAGTC
CAATATTAAAAACTTAGGCACTGACTA_{ACTTTAA}AAATTCTCAA_{ACTATCAATATC}
TAAAGTCATATATT_{TT}TAAGAAAGATTATTCTCAATAACTCTATAAAATAAGTTGAT
GGTTGGCCCATCTAAC_{TTCA}ACTACTATTAGTAAGAAC_{TTTA}ACTTTAATG_{TGTAGTAAG}
GTTTATTCTACCTTTCTCAACATGACACCAACACA_{AT}CAAAACGAAGTTAGT_{GAGGTGC}
TAACATGTGAGGATTAATCCAGT_{GATTCCGGT}CACAATG_{CATTCCAGGAGGAGGTACCCATG}
TCAGTGGAAATTGGCGATATGGTTATT_{TTCTCCCTGATT}GGATAACCAAA_{ATGGAAACA}
GGAGGAGGATAGT_{GATTCTGATGGCATT}CCCTCGATACATT_{CCTGGCTTTCTGGCAA}
AGGGTGCCACATT_{GGAAAGAGGTGAA}ATATAAGTTCTGAA_{ACTGTAGGGAAGAGAACACAT}
TAAGTTAAT_{CAAAGAAAAATCATCATCTATGTTCCAGATTCTCATTAAAGACAAAGTT}
ACCCACAA_{ACTGAGATCACATCTAAGTGCACACTCCTATTGTCAGGTCTAA}ATACATTAAAA
ACCTCATGTGTAATAGGCGTATAATGTATAACAGGT_{GACCAATGTTCTGAATGCATAAAG}
AAATGAATAAAACTCAAACACAGTACTTCTAAACA_{ACTTCAACCA}AAAAAGACCAAAACATG
GAACGAAT_{GGAAAGCTGTAAGGACATGCTTGT}TTAGTCCAGTGGTTCCACAGCTGGCTAA
GCCAGGAGTC_{ACTTGAGGCTT}AAATACAAACATTGGAGCTGGAGGCCATTATCCTTAG
CAAAC_{TAATG}CAGAAACAGAAA_{ACTACCGCATGTTCTCACTT}ATAAGT_{GGGAGGTAAT}
GATAAGAAC_{TTATG}AACACAAAGAAGGAAACA_{ATAGACATTGGAGTCATTGAGAGGGAG}
GGTGGGAGAAGGAAAAGGAGCAGAAAAGATAACTATTGAGTACTGCCTTCACACCTGGGTGA
TGAAATAATATGTACAACAAATCCCTGTGACACATGTTACCTATGGAA_{ACAAACCTTCATGT}
GTATCCCTAAAC_{CTAA}AAAT_{AAAGTT}AAAAAA_{AAAAAA}AAAAAA_{AAAAAA}AAAAAA_{AAAAAA}
AAAAAA_{AAAAAA}AAAAAA_{AAAAAA}AAAAAA_{AAAAAA}AAAAAA_{AAAAAA}AAAAAA_{AAAAAA}

218/246

FIGURE 216

```
></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA82361
><subunit 1 of 1, 352 aa, 1 stop
><MW: 38938, pI: 7.86, NX(S/T): 3
MALLLCFVLLCGVVDFARSLSIITPEEMIEKAKGETAYLPCKFTLSPEDQGPLDIEWLISPA
DNQKVDQVIIILYSGDKIYDDYPDLKGRVHFTSNDLKSGDASINVNLQLSDIGTYQCKVKK
APGVANKKIHLVVLVKPSGARCYVDGSEEIGSDFKIKCEPKEGSLPLQYEWQKLSDSQKMPT
SWLAEMTSSVISVKNASSEYSGTYSCTVRNRVGSQCLLRLNVVPPSNKAGLIAGAIIGTLL
ALALIGLIIFCCRKKRREEKYEKEVHHDIREDVPPPRTSTARSYIGSNHSSLGSMSPSNM
EGYSKTQYNQVPSEDFERTPQSPTLPPAKFKYPYKTDGITVV
```

Signal sequence.
amino acids 1-19

Transmembrane domain:
amino acids 236-257

N-glycosylation sites.
amino acids 106-110, 201-205, 298-302

Tyrosine kinase phosphorylation sites.
amino acids 31-39, 78-85, 262-270

N-myristoylation sites.
amino acids 116-122, 208-214, 219-225, 237-243, 241-247,
245-251, 296-302

Myelin P0 protein.
amino acids 96-125

219/246

FIGURE 217

GATGGCGCAGCCACAGCTTCTGTGAGATCGATTCTCCCCAGTTCCCTGTGGGTCTGAGG
GGACCAAGGGTGAGCTACGTTGGCTTCTGGAAGGGGAGGCATATGCGTCAATTCCCCA
AAACAAGTTGACATTCCCTGAAATGCTATTCTATCTATTCACTGCAAGTGCCTGCT
GTTCCAGGCCTTACCTGCTGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGC
CACGACCTGTGCCACCAACTCGCACTCAGACTCTGAACCTCAGACCTGAAATCTCTTCAC
GGGAGGCTTGGCAGTTTCTTACTCCTGTGGCTCCAGATTTCAGGCCAAGATGAAAGCC
TCTAGTCTTGCCTTCAGCCTCTCTGCTGCGTTTATCTCCTATGGACTCCTCCACTGG
ACTGAAGACACTCAATTGGGAAGCTGTGATGCCACAAACCTTCAGGAATACGAAATG
GATTTCTGAGATAAGGGGCAGTGTGCAAGCCAAGATGGAAACATTGACATCAGAATCTTA
AGGAGGACTGAGTCTTGCAAGACACAAAGCCTGCGAATCGATGCTGCCCTGCGCCATT
GCTAAGACTCTATCTGGACAGGGTATTAAAAACTACCAGACCCCTGACCATTATACTCTCC
GGAAGATCAGCAGCCTCGCCAATTCTTCTTACCATCAAGAAGGACTCCGGCTCTCAT
GCCACATGACATGCCATTGTGGGAGGAAGCAATGAAGAAATACAGCCAGATTCTGAGTCA
CTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGCTTGGGGAACTAGACATTCTC
TGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGATGCTGCTGCTAAGAATATTGAGGT
CAAGAGCTCCAGTCTCAATACCTGCAGAGGAGGCATGACCCAAACCACCATCTCTTACT
GTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTGCAATTGCTTGTGCTGATGAT
TGTCTTATGCATCCCCAATCTAATTGAGACCATACTTGATAAGATTTGTAATATCTT
TCTGCTATTGGATATATTATTAGTTAATATATTATTATTGCTATTAAATGTATT
ATTTTTTACTTGGACATGAAACTTTAAAAAAATTCAAGATTATTTATAACCTGACTAG
AGCAGGTGATGTATTGATACAGTAAAAAAAAACCTTGTAAATTCTAGAAAGAGTGGCT
AGGGGGGTTATTCAATTGACTTCAACTAAGGACATATTACTCATGCTGATGCTCTGTGAGA
TATTGAAATTGAACCAATGACTACTTAGGATGGGTTGTGGAATAAGTTGATGTGGAATT
GCACATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATATTGTGAT
CTTCCAGCCAGGAATCCTACACGGCCAGCATGTATTCTACAAATAAAGTTCTTGCATA
CCAAAAAA

220/246

FIGURE 218

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA83500

MKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNGFSEIRGSVQAKDGN
IDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKISSLANSFLT
IKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEETE

221/246

FIGURE 219

CGCGGAGCCCTGCGCTGGGAGGTGCACGGTGTGCACGCTGGACTGGACCCCCATGCAACCC
GCGCCCTGCGCCTTAACCAGGACTGCTCCGCGCCCCCTGAGCCTCAGGCTCCGGCCGGAC
CTGCAGCCTCCCAGGTGGCTGGGAAGAACTCTCCAACAATAATACATTGATAAGAAAG**AT**
GGCTTAAAAGTGTACTAGAACAAAGAGAAAACGTTTTCACTCTTTAGTATTACTAGGCT
ATTGTGATGTAAAGTGAATTGTGAATCAGGAGACTGTAGACAGCAAGAATTCAAGGATCGG
TCTGGAAACTGTGTTCCCTGCAACCAGTGTGGGCCAGGCATGGAGTTGTCTAAGGAATGTGG
CTTCGGCTATGGGGAGGATGCACAGTGTGACGTGCCGGCTGCACAGGTTCAAGGAGGACT
GGGGCTTCCAGAAATGCAAGCCCTGTCTGGACTGCGCAGTGGTAACCGCTTCAGAAGGCA
AATTGTTCAGCCACCAGTGTGACGCCATCTGCGGGACTGCTTGCCAGGATTTATAGGAAGAC
GAAACTTGTGGCTTCAAGACATGGAGTGTGCTGGAGACCCCTCCTCCTTACG
AACCGCACTGTGCCAGCAAGGTAACCTCGTGAAGATCGCTCCACGGCCTCCAGCCCACGG
GACACGGCGCTGGCTGCCATTCTGAGCGCTCTGCCACCGTCTGCTGGCCCTGCTCAT
CCTCTGTGTCATCTATTGTAAGAGACAGTTATGGAGAAGAAACCCAGCTGGTCTGCGGT
CGCAGGACATTCACTGAGCTACAACGGCTGAGCTGCTGGTACAGACACTCAGCTCCACGAA
TATGCCACAGAGCCTGCTGCCAGTGGCCCTGACTCAGTGCAGACACTGCCGGTGGCG
CTTGCTCCCATCCATGTGCTGTGAGGAGGCCCTGCAGCCCCAACCCGGCGACTCTGGTTGTG
GGGTGCATTCTGCAGCCAGTCTCAGGCAAGAAACGCAGGCCAGCCGGGAGATGGTGCCG
ACTTCTCGGATCCCTCACGAGTCCATCTGTGGCAGTTTCAGATGCCTGGCCTCTGAT
GCAGAATCCCAGGGGGTGACAACATCTTTGTACTCTTATCTGAACACTCACTGGAG
AAGACATTCTCAATCCAGAACTGAAAGCTAACGTCTTGAGATTCAAATAGCAGT
CAAGATTGGTGGGGCTGTTCCAGTCCAGTCTCATTCTGAAAACTTACAGCAGCTAC
TGATTTATCTAGATATAACACACTGGTAGAATCAGCATCAACTCAGGATGCACTAACTA
TGAGAAGCCAGCTAGATCAGGAGAGTGGCGCTGTCATCCACCCAGCCACTCAGACGTCCCTC
CAGGAAGCT**TAA**AGAACCTGCTCTTCTGAGTAGAAGCGTGTGGAAACCCAAAGAGTA
CTCCTTGTTAGGCTATGGACTGAGCAGTCTGGACCTGTCATGGCTCTGGGGCAAAATA
AATCTGAACCAAACGACGGCATTGAAAGCCTTCAGCAGTTGCTCTGAGCCAGACCAGC
TGTAAGCTGAAACCTCAATGAAATAACAAGAAAAGACTCCAGGCCACTCATGATACTCTGCA
TCTTCCTACATGAGAAGCTCTGAGGAGACACTGAAAGACTGATGGTTGAGCT
GGCAGCCTATGAGATTGTGGACATATAACAAGAAACAGAAATGCCCTATGCTTATTTCT
GGTATTGTGGTTTACAAGACTGAAGACCCAGAGTATACTTTCTTCAGAAATAATT
CATACCGCCTATGAAATATCAGATAAATTACCTTAGCTTTATGTAAGATGGGTTCAAAGT
GAGTGTCTATTGAGAAGGACACTTTCTCATCATCTAAACTGATTGTCATAGGTGGTAG
AATGGCCCTCATATTGCTGCCATAATCTGGTTATTAGATGAAGTTACTGAATCAGAG
GAATCAGACAGAGGAGGATAGCTCTTCAGAATCCACACTCTGACCTCAGCCTCGGTCTC
ATGAACACCCGCTGATCTCAGGAGAACACCTGGCTAGGAATGTGGCGAGAAAGGGCAGC
CCATTGCCAGAATTACACATATTGTAGAGACTGTATGCAAAGGTTGGCATATTATATG
AAAATTAGTTGCTATAGAACATTGTTGCATCTGTCCTCTGCTGAGCTTAGAAGGTTAT
AGAAAAAGGGTATTATAAACATAATGACCTTTACTTGCAATTGTAAGATTCTTAAAGGC
TTAGAAATTACACATATCAGGTTCCCTACTACTGAAGTAGCCTCCGTGAGAACACACC
ACATGTTAGGACTAGAACAGAAAATGCACAATTGTAGGGTTGGATGAAGCAGCTGTAACG
CCCTAGTGTAGTTGACCGAGGACATTGTCGTGCTCTCCATTGTAAGATTAGTTAGCA
CATCATCTCCTACTTGTGACCTCCGGTGGATTAAAGAGGACGGTGCTTCTTCTATTAA
AGTGCTCCATCCCTACCATCTACACATTAGCATTGTCAGAGCTAACAGAACAAATTAA
CCCGTTCACTGACAAAGCAGGGAAATGGTCATTACTCTTAATCTTATGCCCTGGAGAAGA
CCTACTTGAAACAGGGCATATTTTAGACTCTGAACATCAGTATGTTGAGGGTACTATGA
TATTTGGTTGGAAATTGCCCTGCCAAGTCAGTCTTTAACTTTAAACTGAATATTAA
AATGTATCTGTCTTCT

222/246

FIGURE 220

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA84210
><subunit 1 of 1, 417 aa, 1 stop
><MW: 45305, pI: 5.12, NX(S/T): 6
MALKVLLQEKTFFTLLVLLGYLSCKVTCESGDCRQQEFRDRSGNCVPCNQCGPGMELSK
ECGFGYGEDAQCVTCLRHLRFKEDWGFQCKPCLDCAVVNRFQKANCSATSDAICGDCLPG
FYRKTKLGVFQDMECVPCGDPPPYEPHCASKVNLVKIASTASSPRDTALAAVICSALAT
VLLALLILCIVYCKRQFMEKKPSWSLRSQDIQYNGSELSCFDRPQLHEYAHRACCQCRD
SVQTCGPVRLLPSMCCEEACSPNPATLGGCVHSAASLQARNAGPAGEMVPTFFGSLTQSI
CGEFSDAWPLMQNPMGGDNISFCDSYPELTGEDIHSLNPELESSTSLSNNSQDLVGGAV
PVQSHSENFTAATDLSRYNNNTLVESASTQDALTMRSQLDQESGAVIHPATQTSIQLQEA
```

Important features of the protein:**Signal peptide:**

Amino acids

1-25

Transmembrane domain:

Amino acids

169-192

N-glycosylation sites:

Amino acids

105-109;214-218;319-323;350-354;368-372;379-383

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

Amino acids

200-204;238-242

Tyrosine kinase phosphorylation site:

Amino acids

207-214

N-myristoylation sites:

Amino acids

55-61;215-221;270-276

Prokaryotic membrane lipoprotein lipid attachment site:

Amino acids

259-270

TNFR/NGFR family cysteine-rich region proteins:

Amino acids

89-96

223/246

FIGURE 221

CTAGAGAGTATAGGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTTAATCCAGGATCCTGTCCTCCTGTCTGAGGAGTGCCTGTCAGTGTGGGGTGAGACAGTTGTCCCACAGGGCTGTCTGAGCAGATAAGATTAAGGGCTGGGTCTGTGCTCAATTAACTCCTGTGGGCACGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCAATGCTGGCCTGCCGTGGAAGGGAGGTCTGTCCTGGCGCTGCTGCTTCTTAGGCTCCAGATCCTGCTGATCTATGCCTGGCATTCCACAGGAGCAAAGGGACTGTGATGAACACAATGTATGGCTCGTTACCTCCCTGCCACAGTGGAGTTGCTGTCACACATTCAACCAACAGAGCAAGGACTACTATGCCTACAGACTGGGCACATCTGAATTCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCTGGGAGAACTAGGTGTGGAAATTGAAGACGACATTGACAACTGCCATTCCAAGAACAGACAGAGCTGAACAATACTTTCACCTGCTTCTTCACCATCAGCACCAGGCCCTGGATGACTCAGTTCACTGCCTCTGAACAAGACCTGCTGGAGGGATTCCACTGAGTGAAACCCACTCACAGGCTTGTCCATGTGCTGCTCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTTCAGTGGCTGAGCAGCTTGGACTTGTATCCTATTGTCATGTGTTGAGATCTCAGATCAGTGTGTTAGAAAATCCACACATCTGAGCCTAATCATGTAAGTAGATCATTAAACATCAGCATTAAAGAAAAAAAAAAAAAAAAAAAAAA

224/246

FIGURE 222

MLGLPWKGGLSWALLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDYYAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFTISTRPWMTQFSLLNKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

225/246

FIGURE 223

AATCGGCTGATTCTGCATCTGGAAACTGCCTTCATCTTGAAAGAAAAGCTCCAGGTCCCT
TCTCCAGCCACCCAGCCCCAAGAGTGGTATGCTGCTGCTGCTTCCGACTGGCTGG
CCTCTCGGTGGCAGAGGGACAAGCATTTCATCTGGGAAGTGGTACGAAATTGAGAAGATCCC
GCAGGAGAATTGACGTGAATAAGTATCTCGGAAGATGGTACGAAATTGAGAAGATCCC
AACAAACCTTGAGAATGGACGCTGCATCCAGGCCAACTACTCACTAATGAAAACGGAAA
GATCAAAGTGTAAACCAAGGAGTTGAGAGCTGATGGAAGTGTGAATCAAATCGAAGGTGA
AGCCACCCCCAGTTAACCTCACAGAGCCTGCCAAGCTGGAAGTTAAGTTCTGGTTAT
GCCATCGGCACCGTACTGGATCCTGGCCACCGACTATGAGAACTATGCCCTCGTGTATT
CTGTACCTGCATCATCCAACCTTTCACGTGGATTTGCTGGATCTTGGCAAGAAACCC
TAATCTCCCTCCAGAAACAGTGGACTCTCTAAAAAAATCCTGACTTCTAATAACATTGA
TGTCAAGAAAATGACGGTCACAGACCAGGTGAACCTGCCCAAGCTCTCGTAACCAGGTT
TACAGGGAGGCTGCACCCACTCCATGTTACTCTGCTTCGCTTCCCTACCCCCACCCCC
CCCCCATAAAGACAAACCAATCAACCACGACAAAGGAAGTTGACCTGAACATGTAACCAT
GCCCTACCCCTGTTACCTTGCTAGCTGCAAAATAAAACTTGTGCTGACCTGCTGTGCTCGC
AAAAAA

226/246

FIGURE 224

MVMLLLSALAGLFGAAEGQAFHLGKCPNPPQENFDVNKYLGRWYEIEKIPPTFENG
RCIQANYSLMENGKIKVLNQELRADGTVNQIEGEATPVNLTEPAKLEVKFWSWMPMSAPY
WILATDYENYALVYSCTCIIQLFHVDFAWILARNPNLPPETVDSLKNILTSNNIDVKKM
TVTDQVNCPKLS

Signal sequence
1-16

N-glycosylation site.

65-68
98-101

cAMP- and cGMP-dependent protein kinase phosphorylation site.

175-178

N-myristoylation site.

13-18
16-21

Lipocalin proteins.

36-47
120-130

Lipocalin / cytosolic fatty-acid binding proteins

41-185

227/246

FIGURE 225

GGGTGATTGAAC TAAACCTCGCCGCACCGAGTTGCAGTACGGCCGTACCCGCACCGCTG
CCTGCTTGC GGTTGGAGAAATCAAGGCCCTACCGGGCCTCCGTAGTCACCTCTATAGTGG
GCGTGGCCGAGGCCGGGGTGACCCCTGCCGGAGCCTCCGCTGCCAGCGACATGTTCAAGGTAA
TTCAGAGGTCCGTGGGGCCAGCCAGCCTGAGCTTGCTCACCTCAAAGTCTATGCAGCACCA
AAAAAGGACTCACCTCCAAAATTCCGTGAAGGTTGATGAGCTTCACTCTACTCAGTCC
TGAGGGTCAATCGAAGTATGTGGAGGAGGCAGCTGAAGAAAGCATCTCACAGC
TCCGACACTATTGCGAGCCATACACAACCTGGTGTCAAGGAAACGTACTCCAAACTAACGCCC
AAGATGCAAAGTTGGTTCAATGGGGTTAGACAGCTATGACTATCTCCAAAATGCACCTCC
TGGATTTTCCGAGACTTGGTTATTGGTTTGCTGGCCTTATTGACTCCTTTGGCTA
GAGGTTCAAAAATAAAGAAGCTAGTGTATCCGCCTGGTTCATGGGATTAGCTGCCTCCCTC
TATTATCCACAACAAGCCATCGTGTGCCCCAGGTCA GTGGGGAGAGATTATGACTGGGG
TTTACGAGGATATATAGTCATAGAAGATTGTGGAAGGAGAACTTCAAAAGCCAGGAAATG
TGAAGAATTCA CCTGGAAC TAAGTAGAAA ACTCCATGCTGCCATCTTAATCAGTTAAGG
TAAACATTGGAAACTCCATAGAATAATCAGTATTCTACAGAAAATGGCATAGAAGTCAG
TATTGAATGTATTAAATTGGCTTCTTCTTCAGGAAA ACTAGACCA GACCTCTGTTATCTT
CTGTGAAATCATCCTACAAGCAA ACTAACCTGGATCCCTCACCTAGAGATAATGTACAAG
CCTAGAACTCCTCATTCTCATGTTGCTATTATGTACCTAATTAAAACCCAGTTAAAAA
AAAAAAAAAAAAAAAAAAAAAAA

228/246

FIGURE 226

MFKVIQRSVGPASLSLLTFKVYAAPKKDSPPKNSVKVDELSLYSVPEGQSKYVEEARSQLLEE
SISQLRHYCEPYTTWCQETYSQTKPKMQSLVQWGLDSYDYLQNAPPGFFPRLGVIGFAGLIG
LLLARGSKIKKLVYPPGFMGLAASLYYPQQAIVFAQVSGERLYDWGLRGYIVIEDLWKENFQ
KPGNVKNSPGTK

Important features:

Signal peptide:

Amino acids 1-23

Transmembrane domain:

Amino acids 111-130

cAMP- and cGMP-dependent protein kinase phosphorylation site:
Amino acids 26-30

Tyrosine kinase phosphorylation site:
Amino acids 36-44

N-myristoylation sites:
Amino acids 124-130;144-150;189-195

229/246

FIGURE 227

CACCGGAGGGCACGCAGCTGACGGAGCTGCCTGCCTCGTTGCCTCGGCCCTCC
ACTGGAGCTGTCGCGCTCCCGCTCCACCGCAGCCCACCCGGCAGAGGAGTCGCTACCA
GCGCCCAGTGCCTCGTCAGTCCGCAAACCTCTGCCGCCGCCGGCTGGGCACCAAA
TACCAGGCTACC**ATGGT**CTACAAGACTCTTCGCTCTTGCACTTAACTGCAGGATGGAG
GGTACAGAGTCTGCCATACATCAGCTCCTTGTCTGTTCTCTCCGACAAACATTGTACAC
CGACCACCATCTGGACTAGCTCTCCACAAAACACTGATGCAGACACTGCCTCCCCATCCAAC
GGCACTCACAACAACCTCGGTGCTCCAGTTACAGCATCAGCCCCAACATCTGCTTCTAA
GAACATTCCATAGAGTCCAGAGAAGAGGAGATCACCAGCCCAGGTTCGAATTGGGAAGGCA
CAAACACAGACCCCTCACCTCTGGTTCTCGTCAACAAGCGGTGGAGTCCACTTAACAACC
ACGTTGGAGGAACACAGCTCGGGCACTCCTGAAGCAGGCGTGGCAGCTACACTGTCGAGTC
CGCTGCTGAGCCTCCACACTCATCTCCCTCAAGCTCCAGCCTCATCACCTCATCCCTAT
CAACCTCACCAACCTGAGGTCTTCTGCTCCGTACTACCAACCATAGCTCCACTGTGACC
AGCACCCAACCCACTGGAGCTCCAACCTGCCACAGCTGAGCCAGTGCCCCAGGAGAAAACACCCCAACAA
CTGTGTCAGGCAAAGTGATGTGAGCTCATAGACATGGAGACCACCACCTTCCCAGG
GTGATCATGCAGGAAGTAGAACATGCATTAAGTTAGGCAGCATCGCCGCAATTACCGTGAC
AGTCATTGCCGTGGTGTGCTGGTGTGGAGTTGCAGCCTACCTAAAAATCAGGCATTCT
CCTATGGAAGACTTTGGACGACCATGACTACGGGTCTGGGGAAACTACAACAACCCCTCTG
TACGATGACTCCT**AA**CAATGGAATATGGCCTGGGATGAGGATTAACGTGTTCTTATTATAA
GTGCTTATCCAGTAGAATTAATAAGTACCTGATGCGCATTGAACGACAATCTTAAGCCCTGT
TTTGTGATGGTTTTGTTCTCCCTCTGGCTGCTACAACCTCCCTTTC
TGGTACAAGAAGAACATTCTTAAAGGTGAGTGGAGGCTGATTGCAAGCTGAAGTGGCCA
GCCTTGACCCAGGCCAGGCCAGACACCACGGTGAAGGCTCTTCCCCACTGCAGGACCCAC
TTTGAGAAGGATCGAGGAGGAGGATTGGGTGTTGTTAGGGTTACTTCAGGGAAACA
TTTCATTTGTTATTCTTAAACTCTATTAGGAAATTACATTAAGTATTAAATGAGGGGA
AAGGAAATGAGCTACGAGGATTCACCTGCATGGAGAGAGCAGGGTTTCAGATT
CTTTTAATCTCTATTATCTGGTGTGCTGACAGGATGCTGCCTGCTGGCTCTACGAGC
TGAAAGCAGCTCTAGCTGCCTAATTATGAAAGATGAAAATAGGAAGTGCCTGGAGGG
GCCAGCAGGTACGGGCAGAATCTCAGGTTGCTGTGGGATCTCAGTGTGCCCTACCT
GTTCTCCCTCCAGGCCACCTGTCTGTAAAGGATGTCTGCTCTGTTCAAAGGCAGCTGG
GATCCCAGCCCACAAGTGATCAGCAGAGTTGCATTCCAAAGAAAAGGCTATGAGATGAGC
TGAGTTATAGAGAGAAAGGGAGGGCATGTACGGTGTGGGAAGTGGAAAGAGAAAGCTGGCG
GGGAGAAGGAGGCTAACCTGCACTGAGTACTCATTAGGACAAGTGAGAATCAGCTATTGAT
AATGGCCAGAGATATCCACAGCTGGAGGCCAGAGACTGTTGCTTATACCCACACAG
CAACTGGTCCACTGCTTACTGTCGTTGATAATGGCTGAAAATGTTAAAAC

230/246

FIGURE 228

MYKTLFALCILTAGWRVQSLPLTSAPLSVSLPTNIVPPTIWTSSPQNTDADTASPSNGTHN
NSVLPVTASAPTSLLPKNISIESREEEITSPGSNWEGTNTDPSPSGFSSTSGGVHLTTLEE
HSSGTPEAGVAATLSQSAAEPPTLISPQAPASSPSSLSTSPEVFSASVTTNHSTVTSTQP
TGAPTAPESPTEESSSDHTPTSHATAEPVPQEKTPTTVSGKVMCELIIDMETTTFPRVIMQ
EVEHALSSGSIAAITVTVIAVVLLVFGVAAYLKIRHSSYGRLLDDHDYGSWGNYNPPLYDDS

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 258-278

N-glycosylation sites.

amino acids 58-61, 62-65, 80-83, 176-179

Casein kinase II phosphorylation sites.

amino acids 49-52, 85-88, 95-98, 100-103, 120-123, 121-124, 141-144, 164-167, 191-194, 195-198, 200-203

Tyrosine kinase phosphorylation site.

amino acids 289-296

N-myristoylation sites.

amino acids 59-64, 115-120, 128-133, 133-138, 257-262, 297-302

231/246

FIGURE 229

CTCCTGCACTAGGCTCTCAGCCAGGGATGATGCGCTGCTGCCGCCGCTGCTGCTGCCGG
CAACCACCCATGCCCTGAGGCCGTTGCTGCTGCCCTCGTCCTTTACCTCCCCTGGC
AGCAGCTGCAGCAGGGCCAAACCGATGTGACACCATATACCAGGGCTCGCCGAGTGTCTCA
TCCGCTTGGGGGACAGCATGGGCCGCGGAGGCAGCTGGAGACCATCTGCAGGTCTTGGAAAT
GACTTCCATGCCTGTGCCTCTCAGGTCTGTCAGGCTGTCCGGAGGAGGCAGCTGCAGTGTG
GGAATCACTACAGCAAGAAGCTGCCAGGCCCGTCCGAATAACTTGCACACTCTGTGCG
GTGCCCGGTGCATGTTGGGAGCGCGGCACAGGCTCCGAACCAACCAGGAGACGCTGCCGG
GCTACAGCGCCTGCACTCCCCATGCCCTGCCCGACTGCTGGCGCTGCTGGCTCTG
GCCCTACCTCCTGAGGCCCTGCCTAGCTTGTGGTTGGGTAGCAGCGCCGTACCTCCAG
CCCTGCTCTGGCGGTGGTGTCCAGGCTCTGCAGAGCGCAGCAGGGCTTTCATTAAAGGTA
TTTATATTGTA

232/246

FIGURE 230

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA92265
><subunit 1 of 1, 165 aa, 1 stop
><MW: 17786, pI: 8.43, NX(S/T): 0
MMRCCRRRCCCCRQPPHALRPLLLLPLVLLPPLAAAAAGPNRCDTIYQGFAECLIRLGDSM
GRGGELETICRSWNDFHACASQVLSGCPEAAAVWESLQQEARQAPRPNNLHTLCGAPVH
VRERGTGSETNQETLRATAPALPMAPAPPLAALALAYLLRPLA
```

Important features of the protein:

Signal peptide:

Amino acids 1-35

Transmembrane domain:

Amino acids 141-157

N-myristoylation site:

Amino acids 127-133

Prokaryotic membrane lipoprotein lipid attachment site:

Amino acids 77-88

233/246

FIGURE 231

AAGTACTTGTGTCCGGGTGGACTGGATTAGCTGCGGAGCCCTGGAAGCTGCCTGTCCTT
CTCCCTGTGCTTAACCAGAGGTGCCCAATGGGTTGGACAATGAGGCTGGTCACAGCAGCACTG
TTACTGGGTCTCATGATGGTGGTCACTGGAGACGAGGATGAGAACAGCCGTGCCCCATGA
GCCCTCTTGGACGAGGACACCCCTTTTGCCAGGGCCTGGAAGTTCTACCCAGAGTTGG
GGAACATTGGCTGCAAGGTTCTGATTGTAACAACACTACAGACAGAACAGATCACCTCCTGG
ATGGAGCCGATAGTCAAGTTCCGGGGCCGTGGACGGCGCAACCTATATCCTGGTATGGT
GGATCCAGATGCCCTAGCAGAGCAGAACCCAGACAGAGATTCTGGAGACATTGGCTGGTAA
CAGATATCAAGGGCGCCGACCTGAAGAAAGGAAGATTCAAGGGCCAGGGAGTTATCAGCCTAC
CAGGCTCCCTCCCCACCGGCACACAGTGGCTTCCATCGCTACCAAGTTCTTGTCTATCTTCA
GGAAGGAAAAGTCATCTCTCCCTCCAAAGGAAAACAAAACCTCGAGGCTCTGGAAAATGG
ACAGATTTCTGAACCGCTTCCACCTGGCGAACCTGAAGCAAGCACCAGTTCATGACCCAG
AACTACCAGGACTCACCAACCCCTCCAGGCTCCCAGAGGAAGGGCCAGCGAGCCCAAGCACAA
AACCAAGGCAGAGATAGCTGCCTGCTAGATAGCCGGCTTGCCATCCGGCATGTGGCCACAC
TGCTCACCAACCGACGATGTGGGTATGGAACCCCTCTGGATACAGAACCCCTTCTTTCCAA
ATTAAAAAAAATCATCAA

234/246

FIGURE 232

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA92274
><subunit 1 of 1, 223 aa, 1 stop
><MW: 25402, pI: 8.14, NX(S/T): 1
MGWTMRLVTAALLGLMMVVTGDEDENSPCAHEALLDEDTLFCQGLEVFYPELGNIGCKVVP
DCNNYRQKITSWMEPIVKFPGAVDGATYILVMVDPDAPSRAEPRQRFWRHVLVTDIKGADLK
KGKIQGQELSAYQAPSPPAHSGFHRYQFFVYLQEKGKVISLLPKENKTRGSWKMDRFLNRFHL
GEPEASTQFMTQNYQDSPTLQAPRGRASEPKHKTRQR
```

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 169-173

Tyrosine kinase phosphorylation site.

amino acids 59-68

N-myristoylation sites.

amino acids 54-60, 83-89, 130-136

Phosphatidylethanolamine signature.

amino acids 113-157

235/246

FIGURE 233

AAGGAGCAGCCGCAAGCACCAAGTGAGAGGC**ATGAAGTTACAGTGTGTTCCCTTGGCTC**
CTGGGTACAATACTGATATTGTGCTCAGTAGACAACCACGGCTCAGGAGATGTCTGATTTC
CACAGACATGCACCATAAGAAGAGAGTTCCAAGAAATCAAAGAGCCATCCAAGCTAAGG
ACACCTTCCCAAATGTCACTATCCTGTCCACATTGGAGACTCTGCAGATCATTAAGCCCTTA
GATGTGTGCTGCGTGACCAAGAACCTCTGGCGTCTACGTGGACAGGGTGTCAAGGATCA
TCAGGAGCCAACCCCCAAATCTTGAGAAAATCAGCAGCATTGCCAACTCTTCCTCTACA
TGCAGAAAATCTGCGGCAATGTCAGGAACAGAGGCAGTGTCACTGCAGGCAGGAAGGCCACC
AATGCCACCAGAGTCATCCATGACAACATGATCAGCTGGAGGTCCACGCTGCTGCCATTAA
ATCCCTGGGAGAGCTGACGTCTTCTAGCCTGGATTAATAAGAATCATGAAGTAATGTTCT
CAGCT**TGAT**GACAAGGAACCTGTATAGTGTGATCCAGGGATGAACACCCCCCTGTGCGGTTACT
GTGGGAGACAGCCCACCTTGAAGGGGAAGGGAGATGGGAAGGCCCCCTGCAGCTGAAAGTCC
CACTGGCTGGCCTCAGGCTGTTATTCCGCTTGAATAGGCAAAAGTCTACTGTGGTAT
TTGTAATAAAACTCTATCTGCTGAAAGGGCTGCAGGCCATCCTGGGAGTAAAGGGCTGCCTT
CCCATCTAATTATTGTAAAGTCATATAGTCATGTCTGTGATGTGAGCCAAGTGATATCCT
GTAGTACACATTGTAAGTGGTTTCTGAATAAAATTCCATATTTACCTATGA

236/246

FIGURE 234

```
></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA92282
><subunit 1 of 1, 177 aa, 1 stop
><MW: 20452, pI: 8.00, NX(S/T): 2
MKLQCVSLWLLGTILILCSVNDNHGLRRCLISTDMHHIEESFQEIKRAIQAKDTFPNVTILST
LETLQIIKPLDVCCVTKNLLAFYVDRVFKDHQEPNPKILRKISSIANSFLYMQKTLRQCQEQRQCHCRQEATNATRVIHDNYDQLEVHAAAIKSLGELDVFLAWINKNHEVMFSA
```

Signal sequence:
amino acids 1-18

N-glycosylation sites.
amino acids 56-60, 135-139

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 102-106

N-myristoylation site.
amino acids 24-30

Actinin-type actin-binding domain signature 1.
amino acids 159-169

237/246

FIGURE 235

GCCC GGCGG CTGCC TTGGG TGC TCC CTC GCCC GAC ACC CAG ACC GAC CTT GACC GC
CCAC CTGG CAGG AGC AGG ACAGG AC GGCG GAC GCG GGC **AT**GG CGAG CTCC GGGG CCCTT
TCT CTG CGGG GCC CTG CTAGG CCTT GCCT GAGT GGG CTGG CC GTGG AGGT GAAGG TAC
CCAC AGAG CGC GCTG AGC AC GCCC CTGG GG AAG ACAG CGAG CTG ACCT GCAC CTAC AGC AC
TCGG TGGG AGAC AGCT CGCC CTGG AGT GGAG CTT GTG CAGC CTGG AAAC CCAT CTCT GA
GTCCC ATCCA AT CCTG TACT TCA CCA ATGG CCAT CTG TAT CCA ACT GTT CTA AGT CAA AGC
GGGTCAGC CTG CTT CAGA ACCCCCCC ACAGT GGGG GTGG CCAC ACTG AAA ACTG ACTG AC GTC
CACCC CTCA GATA CTGG AACCTAC CTG CCA AGT CAAC AACC ACCAG ATTT CTAC ACCAA
TGGG TTGGG GCTA ATCA AACCTACT GTG CTGG TTCCCC CAGTA ATCC CTTATG CAGT CAGA
GTGG ACAC AA CCTCT GTGG GAGG CTACT GCA CTGAG ATG CAG CTCT CGAG GGGG GCT CCT
AAGCC AGT GTACA ACTGGG TGCG TCTG GA ACTTT CCTAC ACCTT CTG CAGC ATGGT
TCAAG ATGAG GTG TCTGG CCAG CTCA TT CACCA ACCT CTCC CTG ACCT CCTC GGG CAC CT
ACCG CTGT GTGG CCAC CA ACCAG ATGG CAGT GCAT CCTG TGA GCT GAC CCT CTG TGA C
GAAC CCCT CCAAGG CGAGT GGCG GAGC TCTG ATT GGGG TGCT CCTGG CGT GCT GTT GCT
GTCAG TTG CTG CGT TCTG CTC GGTC AGGTT CCAG AAAG AGAG GGGG AAGA AGCC CAAGG AGA
CAT ATGGGG TAGT GAC CTG CTT CGGG AGG ATGCC ATCG CTG CCTGG ATCT GAG CAC ACT TGT
ATGAGGG CTG ATT CTAGCA AGGGG TTCT GGAA AGAC CTC GCT GCG CAG CAC CGT GAC GAC
CACCA AGT CCA AGC TCC CTATGGT CGT **TGA** CTT CTCC GAT CC CTGAG GG CGGT GAGGGG
AATATCAATAATTAAAGTCTGTGGGTACCC TTNAAAAAAAA

238/246

FIGURE 236

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108760
><subunit 1 of 1, 327 aa, 1 stop
><MW: 34348, pI: 7.88, NX(S/T): 2
MAELPGPFLCGALLGFLCLSGLAVEVKVPTEPLSTPLGKTAELTCTYSTSVGDSFALEWS
FVQPGKPISESHPILYFTNGHLYPTGSKSKRVSLLQNPPPTVGVATLKLTDVHPSDTGTYL
CQVNNPPDFYTNGLGLINLTIVPPSNPLCSQSGQTSGGSTATRCSSSEGAPKPVYNWV
RLGTFPTPSPGSMVQDEVSGQLILTNLSLTSSGTYRCVATNQMGSAASCETLSVTEPSQG
RVAGALIGVLLGVLLSVAAFCLVRFQKERGKKPKETYGGSDLREDAIAPGISEHTCMRA
DSSKGFLERPSSASTVTTTKSKLPMVV
```

Important features of the protein:**Signal peptide:**

Amino acids 1-20

Transmembrane domain:

Amino acids 242-260

N-glycosylation sites:

Amino acids 138-142; 206-210

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 90-94

N-myristoylation sites:

Amino acids 11-17; 117-123; 159-165; 213-219; 224-230; 244-250;
248-254

Amidation site:

Amino acids 270-274

Prokaryotic membrane lipoprotein lipid attachment site:

Amino acids 218-229

239/246

FIGURE 237

GGATGCAGCAGAGAGGGAGCAGCTGGAAGCCGTGGCTGCGCTCTTCCCTCTGCTGGCG
TCCCTTCTCCAGGGTGTATATCGTCTTTCCCTGGAGATTGCGAGATGCCCATG
TCCGAGGTTATGTTGGAGAAAAGATCAAGTGAAATGCACTTCAGTCAGATG
TCACTGACAAGCTTACTATAGACTGGACATATGCCCTCCCAGCAGCAGGCCACACAGTAT
CAATATTCATTATCAGTCTTCAGTACCCAACCACAGCAGGCACATTGGGGATCGGA
TTTCCTGGGTTGGAAATGTATACAAAGGGGATGCATCTATAAGTATAAGCAACCTACCA
TAAAGGACAATGGGACATTCACTGCTGTGAAGAATCCCCAGATGTGCACCATAATA
TTCCCATGACAGAGCTAACAGTCACAGAAAGGGTTTGCAACCATGCTTCCTCTGTGG
CCCTCTTCCATCCTGTCTTGTGCCCTCAGCCGTGGTGGCTCTGCTGCTGGTGA
GAATGGGGAGGAAGGGCTGCTGGGCTGAAGAAGAGGGAGCAGGTCTGGCTATAAGAAGTCAT
CTATTGAGGTTTCCGATGACACTGATCAGGAGGGAGGAAGGGCGTGTATGGCGAGGCTT
GTGTCCGTTGCGCTGAGTGCCTGGATTCAACTATGAAGAGACATATTGATGAAAGTC
TATGACACAAAGAAGAGTCACCTAAAGACAGGAAACATCCCATTCCACTGGCAGCTAAAGC
CTGTCAGAGAAAGTGGAGCTGGCCTGGACCATAGCGATGGACAATCCTGGAGATCATCAG
TAAAGACTTTAGGAACCACTTATTGAATAATGTTCTGTTGTATTATAAACTGT
TCAGGAAGTCTCATAAGAGACTCATGACTTCCCTTCAATGAATTATGCTGTAATTGAA
TGAAGAAATTCTTTCTGAGCA

240/246

FIGURE 238

MQQRGAAGSRGCALFPLLGVLFQGVYIVFSLEIRADAHVRGYVGEKIKLKCTFKSTSD
VTDKLTIDWTYRPPSSSHTVSIFHYQSFQYPTTAGTFRDRISWVGNVYKGDASISISNP
TIKDNGTFSCAVKNPPDVHHNIPMTELTVTERGFGTMLSSVALLSILVFVPSAVVALL
LVRMGRKAAGLKKRSRSGYKKSSIEVSDDTDQEEEEACMARLCVRCAECLSDYEETY

Transmembrane domain

11-30
157-177

N-glycosylation site

123-127

cAMP- and cGMP-dependent protein kinase phosphorylation site

189-193
197-201

Tyrosine kinase phosphorylation site

63-71

N-myristoylation site

5-11
8-14
124-130
153-159

Amidation site

181-185

241/246

FIGURE 239

CAGGCAGGCCCGCGCGCAGGGCCCTGGACCCGCGGCTCCGGGGATGGTGAGCAAGGCCTGCTGCC
TCGTGCTCGCGTCAACCGCAGGAGGATGAAGCTGCTGGCATGCCCTGCTGCCCTACGTCCTCTGTT
TGGGCAACTCGTTAATATGAGGTCTATCCAGAAAATGGTAACAAAAATTGAAAGCAAGATTGAAGAGAT
GGTGAACCACTAAGAGAGAAAATCAGAGATTAGAAAAAGCTTACCCAGAAATACCCACCACTAAAGTTT
TATCAGAAAAGGATCGGAAAAGAATTGATAACAGGAGGCGCAGGGTCTGGCTCCCATCTAACTGACAA
CTCATGATGGACGGGCCACGAGGTGACCGTGGTGGACAATTCTCACGGGAGGCCCTACATCGAGGTTGACAG
GATCGGACATGAGAACTTCGAGTTGATTAACACAGCTGGTGGAGCCCTACATCGAGGTTGACAG
ACCATCTGGCATCTCCAGCTCCCAAACATGTATAATCCTATCAAGACATTAAAGACCAATACG
GGGACATTAACATGTTGGGCTGGCAAACGAGTCGGTGCCTCTGGCTCCACATCGGAGGTGA
TGGAGATCCTGAAGTCCACCCCTAAAGTGAGGATTACTGGGCCACGTGAATCCAATAGGACCTGGGCTGCT
ACGATGAAGGAAACGTGTTGCAGAGACCATGTGCTATGCCATCGAACAGCAGGAAGGCGTGGAAAGTGC
GCCAGAAATCTCAACACCTTGGGCCACGCATGCACATGAACGATGGCGAGTAGTCAGCAACTTCATCTGCA
GGCCTCCAGGGGAGCCACTACGGTATACGGATCCGGTCTCAGACAAGGGCTCCAGTACGTAGCGATC
TAGTGAATGGCCTCGTGGCTCATGAACAGCAACGTCAAGCAGCCGGTCAACCTGGGAACCCAGAAC
ACAATCCTAGAATTGCTCAGTTAATAAAAACCTTGGTAGCGGAAGTGAAATTCA
CCAGGATGACCCACAGAAAAGAAAACCAGACATAAAAAGCAAAGCTGATGCTGGGTGGGAGCCCGTGGTCC
CGCTGGAGGAAGGTTAAACAAAGCAATTCACTACTTCGTAAGAACCTCGAGTACCCAGGCAAATAATCAGTAC
ATCCCCAAACCAAAGCCTGCCAGAATAAAGAAAGGACGGACTCGCCACAGCTGAACCTCACTTTAGGACAC
AAGACTACCATGTAACATTGATGGGATGTATTTTGGCTTTTTGTTGCTGGTTAAAGAAAGACTTTAAC
GGTGTATGAAGAACAAACTGGAATTCTCATCTGAAGCTGCTTAATGAAATGGATGCTAAAGCTCCCC
TCAAAAAACTGCAGATTTCGCTTGACTTTGAATCTCTTTTATGTAAGAACAGCTGAGATGCATCTG
CGTATTTCAAGTTTTATCTGCTGTGAGAGCATATGTTGACTGCTGTTGACAGTTTATTTACTGTT
TCTTGTGAAGCTGAAAGGAACATTAAGGGACAAAAATGCCGATTTATTTATAAAAGTGGGTACTTAAT
AAATGAGTCGTATACTATGCATAAGAAAATCCTAGCAGTATTGTCAGGTGGTGGCGCCGGCATTGATTT
TAGGGCAGATAAAAGAATTCTGTTGAGAGCTTATGTTCTTTAATTCAAGGTTTCCAAGGTCTACTT
TTGAGTTGCAAACCTGACTTTGAAATATTCTGTTGCTGATCAAGGATATTGAAATCACTACTGTT
GCTCGTATCTGGGCGGGGGCAGGGTGGGGGACAAAGTTAACATATTCTGTTAACATGGTTAAATG
CTATTTAATAAAATATTGAAACTCA

242/246

FIGURE 240

MVKALLRLVSAVNRRRMKLLGIALLAYVASVWGNFVNMRSIQENGELKIESKIEEMVEPL
REKIRDLEKSFTQKYPPVKFLSEKDRKRILITGGAGFVGSHLTDKLMMGDHEVTVDNFFTG
RKRNVEHWIGHENFELINHDVVEPLYIEVDQIYHLASPASPPNYMYNPIKTLKTNTIGTLNM
LGLAKRVRGARLLLASTSEVYGDPEVHPQSEDYWGHVNPIGPRACYDEGKRVVAETMCYAYMKQ
EGVEVRVARIFNTFGPRMHMDGRVVSNFILQALQGEPLTVGSGSQTRAFQYVSDLVNGLV
ALMNSNVSSPVNLGNPEEHTILEFAQLIKNLVGGSEIQLSEAQDDPQKRKPDIKKAKLML
GWEPVVPLEEGLNKAIHYFRKELEYQANNQYIPKPKPARIKKGRTRHS

Important features:

Signal peptide:

amino acids 1-32

N-glycosylation site:

amino acids 316-320

Tyrosine kinase phosphorylation site:

amino acids 235-244

N-myristoylation sites:

amino acids 35-41, 101-107, 383-389

Amidation sites:

amino acids 123-127, 233-237

243/246

FIGURE 241

GCCCGGTGGAGAATTAGGTGCTGCTGGGAGCTCTGCCCTCCCACAGGATTCCAGCTGCAGGG
AGCCTCAGGGACTCTGGGCCGACGGAGTTGGGGCATTCCCCAGAGAGCGTCGCCATGGTC
TGCAGGGAGCAGTTATCAAAGAATCAGGTCAAGTGGGTGTTGCCGCATTACCTGTGTGTC
TGTGGTGGTCATTGCCGAATAGTCCTGCCATCACCTGCCGGCAGGCTGTGAGCTGGCAGATCAGCCGG
AGGCCTGCAGCCCTGATGCCGACATGCTGGACTACCTGCTGAGCCTGGCAGATCAGCCGG
CGAGATGCCCTGGAGGTACCTGGTACCCACGCCAACAGCAAGAAAGCCATGACAGCTGC
CCTGAACAGCAACATCACAGTCCCTGGAGGCTGACGTCAATGTAGAAGGGCTGGCACAGCCA
ATGAGACAGGAGTCCCATATGGCACACCCCCCCTATCTACAGTGACAACACACTGGAG
CACTGGCTGGACGCTGTGCTGGCTCTTCCAAAAGGGCATCAAACGGACTCAAGAACAT
CAAGGCAGTGGGCCCTCCCTGGACCTCTGCCAGCTGACAGAGGAAGGCAAAGTCCGGC
GGCCCATATGGATCAACGCTGACATCTAAAGGGCCCAACATGCTCATCTCAACTGAGGTC
AATGCCACACAGTCCCTGGCCCTGGTCCAGGAGAAGTATCCAAGGCTACCCCTATCTCCAGG
CTGGACCACCTCTACATGTCCACGTCCCCAACAGGACGTACACCCAAGCCATGGTGGAGA
AGATGCACGAGCTGGTGGAGGAGTGCCTCAGAGGGTACCTTCCCTGTACGGTCTTCCATG
GTGCGGGCTGCCTGGCCCACTTCAGCTGGCTGCTGAGCCAATCTGAGAGGTACAGCCTGAC
GCTGTGGCAGGCTGCCTCGGACCCATGTCGGTGAAGATCTGCTCTACGTCCGGGATAACA
CTGCTGTCCACCAAGTCTACTATGACATCTTGAGCCTCTCCTGTACAGTTCAAGCAGCTG
GCCCTGAATGCCACACGGAAACCAATGTACTACACGGGAGGCAGCCTGATCCCTTTCTCCA
GCTGCCTGGGATGACGGTCTGAATGTGGAGTGGCTGGTCTGACGTCCAGGGCAGCGGTA
AAACAGCAACAATGACCCCTCCCAGACACAGAAGGCATGATCCTGCTGAACACTGGCCTCGAG
GGAACGTGGCTGAAAACCCCGTGCCATTGTTCATACTCCAAGTGGCAACATCCTGACGCT
GGAGTCCTGCCTGCAGCAGCTGGCCACACATCCGGACACTGGGGCATCCATTGCAAATAG
TGGAGCCCGCAGCCCTCCGGCCATCCCTGGCCTGCTGGCACGCCCTCCAGCCTTGGCCTC
TTGCATTGGCCTGTGTTGGGTTGGGCAAAATCTCCACGGGAGTTTTCGGTCCCCGGCCA
TGTGGCTGGCAGAGAGCTGCTTACAGCTGTGGCTGAGGTCTCCCCCACGTGACTGTGGCAC
CAGGCTGGCCTGAGGAGGTGCTGGCAGTGGCTACAGGGAACAGCTGCTCACAGATATGCTA
GAGTTGTGCCAGGGCTCTGGCAACCTGTGTCTCCAGATGCAGGCCATGCTGCTGGCCA
CAGCACAGCTGGAGCCATAGGCAGGCTGCTGGCATCCTCCCCCGGGCACCGTCACAGTGGAG
CACAACCCAGCTGGGGCGACTATGCCTCTGTGAGGACAGCATTGCTGGCAGCTAGGGCTGT
GGACAGGACCCGAGTCTACTACAGGCTACCCAGGGTACCAAGGACTTGCTGGCTCATG
TTGGTAGAAACTGAGCACCCAGGGGTGGTGGGCCAGCGGACCTCAGGGCGGAGGCTTCCCAC
GGGGAGGCAGGAAGAAATAAAGGTCTTGGCTTCTCAGGCAAAAAAAAAAAAAAAAG
AAAAAAAAAAAAAAAAAAAAAAAG

244/246

FIGURE 242

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA119514
><subunit 1 of 1, 585 aa, 1 stop
><MW: 64056, pI: 6.58, NX(S/T): 5
MVCREQLSKNQVKWVFAGITCVVVVIAIVLAITLRRPGCELEACSPDADMIDYLLSLG
QISRRDALEVTWYHAANSKKAMTAALNSNITVLEADVNVEGLGTANETGVPIMAHPPTIY
SDNTLEQWLDALVGSSQKGIKLDFKNIKAVGPSLDLLRQLTEEGKVRPPIWINADILKGP
NMLISTEVNATQFLALVQEKPATLSPGWTTFYMSTSPNRTYTQAMVEKMHELVGGVPQ
RVTFPVRSSMVRAAWPHFSWLLSQSERYSLTLWQAASDPMSEDLLYVRDNTAVHQVYYD
IFEPLLSQLKQLALNATRKPMYYTGGSLIPLLQLPGDDGLNVEWLVPDVQGSGKTATMTL
PDTEGMILLNTGLEGTVAENPVPIVHTPSGNILTLESCLQQLATHPGHWGIHLQIVEPAA
LRPSLALLARLSSLGLLHWPVVVGAKISHGSFSVPGHVAGRELLTAVAEVFPHTVAPGW
PEEVLGSGYREQLLTDMLELCQGLWQPVSFQMQAMLLGHSTAGAIGRLLASSPRATVTVE
HNPAGGDYASVRTALLAARAVDRTVYYRLPQGYHKDLLAHVGRN
```

Important features of the protein:**Transmembrane domain:**

Amino acids 18-37 (Possible type II)

N-glycosylation sites:

Amino acids 89-93;106-110;189-193;220-224;315-319

Tyrosine kinase phosphorylation site:

Amino acids 65-74

N-myristoylation sites:

Amino acids 101-107;351-357;372-378;390-396;444-450;545-551

Aminotransferases class-V pyridoxal-phosphate attachment site:

Amino acids 312-330

245/246

FIGURE 243

CTTCAGAACAGGTTCTCCTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCA**ATGGC**
CGCCCTGCAGAAATCTGTGAGCTCTTCCTTATGGGGACCCCTGGCCACCAGCTGCCTCCTTC
TCTTGGCCCTCTGGTACAGGGAGGAGCAGCTGCCCATCAGCTCCACTGCAGGCTTGAC
AAGTCCAACCTCCAGCAGCCCTATATCACCAACCGCACCTCATGCTGGCTAAGGAGGCTAG
CTGGCTGATAACAAACACAGACGTTCTCATTTGGGAGAAACTGTCCACGGAGTCAGTA
TGAGTGAGCGCTGCTATCTGATGAAGCAGGTGCTGAACCTCACCTTGAAGAAGTGCTGTT
CCTCAATCTGATAGGTTCCAGCCTTATATGCAGGAGGTGGTGCCTCTGGCCAGGCTCAG
CAACAGGCTAACGCACATGTCAATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGAAA
AGCTGAAGGACACAGTAAAAAGCTTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTG
GATTTGCTGTTATGTCTCTGAGAAATGCCCTGCATT**TGA**CCAGAGCAAAGCTGAAAATGAA
TAACTAACCCCCCTTCCCTGCTAGAAATAACAATTAGATGCCCAAAGCGATTTTTTAAC
CAAAGGAAGATGGGAAGCCAACCTCCATCATGATGGGTGGATTCCAAATGAACCCCTGCGT
TAGTTACAAAGGAAACCAATGCCACTTTGTTATAAGACCAAGGGTAGACTTCTAAGCA
TAGATATTATTGATAAACATTTCATTGTAACTGGTGTCTATAACAGAAAACAATTATT
TTAAATAATTGTCTTTCCATAAAAAGATTACTTCCATTCTTAGGGAAAAAACCC
CTAAATAGCTTCATGTTCCATAATCAGTACTTTATTTATAATGTATTATTATTATTA
TAAGACTGCATTTATTATCATTATTAAATATGGATTATTATAGAAACATCATTG
ATATTGCTACTTGAGTGTAAAGGCTAATTGATATTGACAATAATTATAGAGCTATAAC
ATGTTATTGACCTCAATAAACACTTGGATATCC

246/246

FIGURE 244

MAALQKS VSSFLMGTLAT SCLLL ALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKE
ASLADNNTDVR LIGEKL FHGVSM SERCYLMKQV LNF TLEEVLF PQSDRF QPYM QEV VPFLAR
LSNRLSTCHIEGDDLHIQRNVQKLKDTVKKLGESEIKAIGEL DLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

INTERNATIONAL SEARCH REPORT

Inte
nt Application No
PCT/US 10/17800

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 B01D39/16

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 B01D

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 4 215 682 A (DAVIS CHARLES I ET AL) 5 August 1980 (1980-08-05) column 6, line 23 – line 31; figure 4 -----	1,2,7,9
Y	US 5 645 627 A (LIFSHUTZ NORMAN ET AL) 8 July 1997 (1997-07-08) Summary of the invention column 4, paragraph 2 – paragraph 3 column 6, paragraph 2 – paragraph 3 -----	1,2,7,9



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the International filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

T later document published after the International filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

& document member of the same patent family

Date of the actual completion of the International search

10 September 2001

Date of mailing of the International search report

18/09/2001

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Plaka, T

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inte
nal Application No
PCT/US 10/17800

Patent document cited in search report	Publication date	Patent family member(s)		Publication date	
US 4215682	A	05-08-1980	AT 380126 B AT 82779 A AU 507773 B AU 4391279 A BR 7900546 A CA 1122546 A CH 642277 A DE 2904170 A DK 45179 A, B, ES 477436 A FR 2416535 A GB 2015253 A, B IT 1116822 B JP 1473168 C JP 54113900 A JP 59000124 B JP 1721198 C JP 2007671 B JP 62290477 A NL 7900855 A NO 790347 A, B, SE 444893 B SE 7900929 A		10-04-1986 15-08-1985 28-02-1980 06-09-1979 11-09-1979 27-04-1982 13-04-1984 09-08-1979 07-08-1979 01-09-1980 31-08-1979 05-09-1979 10-02-1986 27-12-1988 05-09-1979 05-01-1984 24-12-1992 20-02-1990 17-12-1987 08-08-1979 07-08-1979 20-05-1986 07-08-1979
US 5645627	A	08-07-1997	AU 4915496 A WO 9626783 A	18-09-1996 06-09-1996	

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.